THE AMERICAN JOURNAL OF PSYCHOLOGY

EDITED BY

KARL M. DALLENBACH
UNIVERSITY OF TEXAS

AND

M. E. BITTERMAN
BRYN MAWR COLLEGE

E. B. NEWMAN
HARVARD UNIVERSITY

WITH THE COÖPERATION OF

VOL. LXXVIII
MEZES HALL, THE UNIVERSITY OF TEXAS
AUSTIN, TEXAS 78712

TABLE OF CONTENTS

ARTICLES AND NOTES

BOOK REVIEWS

(The reviewer's name appears in parentheses after the title of the work.)

The American Journal of Psychology

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. Stanley Hall

Vol. LXXVIII | MARCH, 1965 | No. 1

SOME NEW STEREOSCOPIC PHENOMENA AND THEIR IMPLICATIONS FOR THE THEORY OF STEREOPSIS*

By Lloyd Kaufman, Sudbury, Massachusetts

Panum's limiting case provides the minimal stimulus-conditions for the occurrence of stereopsis. As shown in Fig. 1, it is comprised of two lines presented to one eye, and one line presented to the other eye. The rule is that, when the single line (*B*) is presented to the right eye and fused with one of the line in *A*, a binocularly perceived pair of lines will be so arranged in depth that the right-hand line is forward and the left-hand line back. This is true when *B* is fused with either line in *A*. When *A* and *B* are interchanged, with *A* presented to the right eye and *B* to the left, then the left-hand line is seen as forward and the right-hand line as back. This, too, is independent of which line in *A* is fused with *B*.

The depth-effect from Panum's limiting case occurs only when the space between the lines in *A* is small. A very weak and unstable effect may occur with separations as great as 2° in central vision, but an optimal or patent effect is best achieved with separations no greater than the size of Panum's fusional area.[1] Separations of about 20′ of arc or less seem suitable for the patent effect.

If the rule stated above is valid, the classical of fusional interpretation of stereopsis can account for Panum's limiting case. This interpretation holds that, when one point in space is fixated, other points at different

* Received for prior publication October 15, 1964 from the Sperry Rand Research Center. This research was partially supported by the U. S. Army Medical Research and Development Command under Contract No. DA-49-193-MD-2654. The author is indebted to Colin Pitblado for his assistance in conducting the experiments, to George Harker for criticisms and advice, and to Shimon Even for helpful discussions of the mathematical concepts of correlation and phase.

[1] K. N. Ogle, Part II: The optical space sense, in Hugh Davson (ed.), *The Eye*, 4, 1962, 375.

distances will be imaged at non-corresponding places in the two eyes. If this disparity is not too great, then the fixated and non-fixated points will be seen at different distances. The direction and magnitude of the difference in distance is some function of the amount and direction, or cross, of the disparity. According to the classical interpretation, the disparate half-images interact, in some way, in the nervous system. This interaction has been termed 'fusion.' The fusion produces the depth-effect.[2]

There is some ambiguity in the use of the term 'fusion.' First, sensory fusion must be distinguished from motor fusion.[3] Motor fusion refers to the tendency of the eyes to converge to bring similar disparate images into

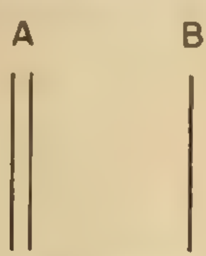

Fig. I. Panum's Limiting Case

binocular register. Sensory fusion refers to the interaction between disparate half-images—images that are not in good binocular register. In this paper we shall mean sensory fusion when using the term 'fusion.'

Even 'sensory fusion' is used ambiguously. One interpretation is that disparate half-image representations actually combine at different depths in a neural net, or other medium, in the brain. Moreover—and this is a crucial distinction—this combining of neural processes is accompanied by a phenomenal fusion of the half-images. According to Koffka, the energy expended in achieving this fusion produces the depth-response.[4] This view is untenable as it stands. It is well known that it is possible to achieve stereopsis with visible half-images.[5] Werner, therefore, makes use of the concept of 'allelotropia,' or the tendency for visible half-images so to shift

[2] Typical accounts of this classical view of stereopsis are given by E. G. Boring (*Sensation and Perception in the History of Experimental Psychology*, 1942, 282-288, 301-303), Armin von Tschermak-Seysenegg (*Introduction to Physiological Optics*, Paul Beeder, transl., 1952, 180-183), and Arthur Linksz, *Physiology of the Eye*, vol. 2, *Vision*, 1952, 380.

[3] Von Tschermak-Sysenegg, *op. cit.*, 182-183.

[4] Kurt Koffka, *Principles of Gestalt Psychology*, 1935, 268-270.

[5] Ogle, Precision and validity of stereoscopic depth from double images. *J. opt. Soc. Amer.*, 43, 1953, 906-913.

their apparent visual directions that they are closer together when viewed binocularly.[6] This is a tendency toward phenomenal fusion which, in Von Tschermak-Seysenegg's system, is the correlate of neural fusion.[7] Phenomenal fusion is not, therefore, the basis for stereopsis, but neural or sensory fusion is. This neural fusion is indicated by allelotropia which may or may not be complete. The phenomenal counterpart of neural fusion is the depth-response. Hence, for the present, it is sufficient to define 'fusion' as some kind of matching and combining of the neural representations of similar point-domains comprising monocular inputs.

Panum's limiting case falls within this general view if the assumption is made that the single line, *B*, in Fig. 1, can fuse with both of the contralateral lines in *A*. Thus, when *B* is in register with one of the contralateral lines, it fuses with it to form a point of reference. Since it fuses also with the second line, with respect to which it is disparate, the second line will be seen at a different distance than the reference-line. This interpretation, which is given by Ronne and by Von Tschermak-Seysenegg, is consistent with Hering's interpretation as stated by Ogle.[8] In this version, single retinal image has the double function of establishing the directional value of the fixation-line and also 'cooperates' with the second image as either a crossed or uncrossed disparate image. This cooperation of the images has been termed 'fusion.'

A phenomenon similar to Panum's limiting case was predicted in the course of a more general study of binocular vision.[9] This new phenomenon, together with certain variations of it, raises some serious questions about the traditional interpretation of the limiting case, and about stereopsis in general.

DEMONSTRATION I

A stereogram capable of producing the basic phenomenon is shown in Fig. 2. It is comprised of two matrices of letters of the alphabet. Both matrices contain identical letter-groupings, 11 letters wide and 4 rows high. In addition, the top row in the right-hand matrix, which contains 13

[6] Heinz Werner (Dynamics in binocular depth perception, *Psychol. Monogr.*, Whole No. 218, 49, 1937) provides evidence for the shift in visual direction of half-images when viewed simultaneously.

[7] Von Tschermak-Seysenegg, *op. cit.*, 180.

[8] Gerhard Ronne, The physiological basis of sensory fusion, *Acta Ophthalmologica*, 34, 1956, 1; Von-Tschermak-Seysenegg, *op. cit.*, 182. Ogle, in H. Davson, *op. cit.*, 374-376.

[9] Lloyd Kaufman, On the spread of suppression and binocular rivalry, *Vision Research*, 3, 1963, 401-415; Suppression and fusion in viewing complex stereograms, this JOURNAL, 77, 1964, 193-205; On the nature of binocular disparity, this JOURNAL, 77, 1964, 393-402.

letters, and the bottom row of the left-hand matrix, have no counterparts in their contralateral fields. This is also true of the left-hand column of the right matrix and the right-hand column of the left matrix. When the patterns are superposed, therefore, there is an inner rectangular array of letters formed by the superpositioning of identical inputs to the two eyes. Without fixation-disparity, this inner array contains no disparate letters. This inner rectangular array is surrounded by a 'ring' of letters,

```
                         qwertyuioplkj
 gcasdewqrdes            fgcasdewqrde
 oiuytgfdsazv            aoiuytgfdsaz
 moiuwqfdsvbo            lmoiuwqfdsvb
 vbnmlkjhgfdb            xvbnmlkjhgfd
asdfghjklpoiu
```

FIG. 2. A STEREOGRAM WITHOUT DISPARITY BETWEEN SIMILAR DISPLAY-ELEMENTS

the left side and top originating in the right eye, and the right side and bottom originating in the left eye. With fixation of the inner letters, a depth-effect will appear. *O* will see the inner array as though through a window in the outer ring. The effect is not unlike that experienced in viewing Julesz's dot-patterns.[10] When the two matrices are interchanged, the rectangular array is seen in front of the outer ring. This, too, is similar in appearance to Julesz's stereo-effect.

EXPERIMENT I

The foregoing description of the effect was confirmed by five sophisticated *O*s. Four of the *O*s had no prior knowledge of the direction of depth, but they reported it correctly. There was however, a considerable delay before some of them obtained any depth-effect at all, and it therefore was considered desirable to conduct the following experiment.

Method. The stimulus used was the same as that shown in Fig. 2. The two half-fields were made into black on white 35-mm. slides and inserted into a *TDC Stereo Project-or-View* device, which caused the two half-field images to be polarized and projected into perfect visual register on a ground-glass screen. The luminous screen with black letters was viewed in a darkened room from a distance of 165 cm. The letter-separation on the screen was about 0.4 cm. center-to-center. *O* placed his chin in a rest and wore polaroid-filter spectacles which gave him a haploscopic view of the pattern. A half-silvered front-surface mirror was mounted at a 45° angle to *O*'s line of sight between his eyes and the screen. Off to one side, at 90°

---

[10] Bela Julesz, Binocular depth perception of computer-generated patterns, *Bell System Tech. J.*, 39, 1960, 1125-1162; B. W. White, Stimulus-conditions affecting a recently discovered stereoscopic effect, this JOURNAL, 75, 1962, 411-420.

to the line of sight, was a small red-light source which was reflected from the mirror into O's visual field. The spot of light could be seen at any plane of depth in the space in front of O, because it was a virtual image and its optical distance was capable of adjustment by moving the source on a meter-stick track toward and away from the combining glass. By instructing E to move the spot, O could place it in front of the stereogram, within it, and even behind it.

O first was shown a conventional stereogram, similar to Julesz's regular pattern, made from randomly selected letters of the alphabet arranged in rectangular matrices.[11] This procedure would have permitted the elimination of any O incapable of achieving stereopsis. It also served to acquaint O with the procedure. O instructed E so to move the spot that it was in the same depth-plane as the elevated (or depressed) inner rectangle and also in the same plane as the background-letters. Then the spot was turned off, and the test-stereogram inserted. O was required simply to look at the stereogram for a few moments. He then was asked if he saw depth. If O saw depth, he was asked to describe what it was that he saw in depth, and what the direction of depth was. O had no prior knowledge of the relationship between the orientation of the polaroid spectacles and the direction of depth. He remained ignorant of this relationship throughout the experiment, since E inverted the spectacles or kept them in the same orientation between each measurement, in a random order. It was decided in advance that if, after several minutes, O did not report depth, he would be prompted. This was to be done at three different levels. First, he would be told that the inner region would be at a different depth than the outer. Secondly, if this led to poor results, he would be told the direction of depth. Third, he would be tested regardless, with the spot being adjusted to the depth of the inner and outer letters.

Each O was given eight ascending and descending trials by the method of average error. Since two measures were made in each direction, one for the furthest and one for the nearest depth-plane, each O gave 16 measures for each orientation of the filters, or 32 measures in all. First, the spot was shown either well in front of the over-all pattern or well behind it. If the spot was in front of the pattern, it would be moved back until it was reported as being in the same plane as the nearer of the letters, the annulus or the inner region, depending upon the orientation of the spectacles. O could correct his adjustments as often as necessary to achieve an acceptable judgment of equal depth. Then the spot-movement was continued in the backward direction until it was seen in the same plane as the more distant array of letters. The reverse of this procedure was used in going from behind the pattern towards O.

Subjects. Five women and four men served in the experiment. All were employees of the Research Center. All reported normal or corrected-to-normal vision.

Results. Four of the Os reported depth and its correct direction spontaneously. Two required prompting for one direction of depth only. Three required prompting for both directions, and only one of these finally demonstrated a depth-response. (His measurements were different from the other six.) The latter three Os gave evidence of a fixation-disparity. They could not see a rectangular over-all matrix, but instead saw over-

[11] Kaufman, on the nature of binocular disparity, 193-205.

hanging letters in the upper right-hand corner and lower left-hand corner of the pattern. It is assumed that depth was lost because of the fixation-disparity. Disparity does not affect depth in conventional stereograms, which explains why they had no trouble with the first pattern they were shown.

The mean depth for the eight *O*s who gave the response when the inner square was seen behind the outer ring of letters was 10.2 cm. When the inner square was seen in front of the outer ring of letters the mean depth was 10.9 cm. Individual data are shown in Table I.

The similarity between the responses of those *O*s who had no fixation-disparity is evident. The 16 forward-responses of each were generally alike, and so also were the 16 back-responses.

TABLE I

MEAN DIFFERENCES BETWEEN PERCEIVED FORWARD AND BACK SETTINGS OF SPOT IN EXPERIMENT I

O	Square behind	Square in front
JM*	7.7	4.6
BLB†	11.5	11.0
MFC*	0	0
JW*	0	0
HM	12.8	14.4
DW	11.2	13.5
DH	6.0	11.5
LAC‡	11.0	10.4
NP	11.1	11.1

* Exhibited evidence of fixation-disparity and were prompted.

† Needed prompting only when inner square should have been seen behind. Responded spontaneously when inner square was in front.

‡ Needed prompting only when inner square should have been seen in front. Responded spontaneously when inner square was behind.

The major difference between this phenomenon and that experienced in most stereograms is that there is no disparity between identical forms any place in the binocular pattern. The left-side column has no obvious disparate counterpart within the matrix, and the right-side column also exists separately. Another curious thing about the pattern is that the top and bottom rows always appear to be in the same plane as the end rows—this in spite of the fact that, as uniocular images, the depth of these rows would usually be in the most distant plane of ordinary stereograms.[12]

It might be argued that this stereoscopic effect is due to the presence of

[12] Julesz, *op. cit.*, 1159. See also Julesz, Binocular depth perception without familiarity cues, *Science*, 45, 1964, 356-362. This generalization is contingent, however, upon the absence of secondary depth-cues; *e.g.* as in Ogle, in Davson, *op. cit.*, 385.

a disparity between the two end-columns. That is, since one end-column is in one eye and the other end-column is in the other eye, they are falling on disparate places while the inner region exhibits no disparity. This disparity between the two end-columns determines the stereopsis. There are several things wrong with this argument. First, the end-columns contain different forms. Secondly, these different forms can be separated by as much as 7° and the effect still is present. Finally, the end-columns both are either nasal or temporal when the inner region is fixated. Their representations may, therefore, terminate in different cerebral hemispheres.

None of these objections is conclusive, of course. Disparity between dissimilar forms has been shown to be capable of producing depth when they are embedded in similar patterns or groups.[13] Perhaps the columnar

FIG. 3. STEREOGRAM SIMILAR TO FIG. 2, BUT WITH ONLY ONE UNIOCULAR COLUMN

arrangements of the forms are themselves the disparate stimuli. Furthermore, recent evidence indicates that there may be a secondary cross-over of fibers from the lateral geniculate bodies through the corpus callosum to mediate the termination of appropriate fibers in common hemispheres.[14] Finally, certain ordinary stereograms also possess temporal-temporal disparities but still yield a depth-effect.[15] Clearly, then, the only satisfactory test of the hypothesis that the depth-effect is due to disparity between the end uniocular columns is to see if it occurs also when one of these columns is absent.

DEMONSTRATION II

A stereogram similar to that of Fig. 2 is shown in Fig. 3. It differs because only one uniocular end-column is present. Viewing this pattern in a stereoscope causes the binocular region to appear behind the left-hand uniocular column. Of course, the inner identical letters must be in

---

[13] Kaufman, *op. cit.*, this JOURNAL, 77, 1964, 396-398.

[14] M. Glickstein, J. Miller, and O. A. Smith, Jr., Lateral geniculate nucleus and cerebral cortex: Evidence for a crossed pathway, *Science*, 145, 1964, 159.

[15] C. E. Osgood (*Method and Theory in Experimental Psychology*, 1953, 258) points out that Werner's displaced half-images fall on nonhomonymous hemiretinas and therefore questions the plausibility of neural fusion or semi-fusion of these stimuli in the visual area (17).

register before the effect can appear. The depth-effect is greatest at the end of the inner array closest to the uniocular column, and appears to diminish with distance away from it. At the opposite end, the inner array appears to be in nearly the same plane as the top and bottom uniocular rows. This impression was confirmed by five *Os* without prior knowledge of the predicted direction of depth. It must, therefore, be concluded that the phenomenon is not due to a disparity between the uniocular rows, but instead is based upon a stimulus-situation analogous to Panum's limiting case, *i.e.* upon the fact that a uniocular column is presented together with at least one fused binocular column.

One way in which this effect differs from the classical limiting case is that more than one binocular row is seen in depth. It might be concluded, therefore, that the uniocular image is 'cooperating' or 'fusing'

```
asdfghjklp                lqwertyui
 qwertyuio                zxcvbnmklu
```

FIG. 4. THE STEREOSCOPIC EFFECT WITH ONE ROW OF BINOCULAR LETTERS

with several different sets of binocular stimuli simultaneously. This means that it must be fusing with several different sets of forms at the same time, even though it would rival with these forms if they were superposed.

Simultaneous cooperation between the uniocular form and several different contralateral members of binocular forms could be mediated only by common properties of these forms. At least this is demanded by the classical interpretation. Here the common property may be the columnar arrangement of the forms, but the columnar arrangement is a stimulus only by hypothesis, because, in the monocular fields, the perceived organization is the arrangement of letters in rows (an effect of proximity). Moreover, the effect is obtainable with but one binocular row of letters, where the letters are identical, and a uniocular ring of letters, as in Fig. 4. Alternatively, therefore, the common property may be a subjective contour surrounding the individual letter, *i.e.* each letter fills, roughly, a rectangular field. If all or most letters will, for example, serve equally well to excite a cortical collector-cell similar to the cells found by Hubel and Weisel then, perhaps, the fusion is between impulses from these cells.[16]

[16] Julesz (*op. cit.*, *Science*, 145, 1964, 360) suggests that depth-perception results from processing a pre-formed binocular field. He cites Hubel and Wiesel as providing neurological evidence which is in agreement with his belief (D. H. Hubel and T. N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's cortex. *J. Physiol.*, 160, 1962, 106-154). This would localize the fused binocular 'scene' in the visual cortex prior to depth-'recognition.' Hubel

Hence, the form within the region filled by the letter is not a revelant stimulus, but the shape of the region occupied by the letter is.

DEMONSTRATION III

If it is true that the visual system is responding in this case to regions of excitation having a roughly similar size and shape, independently of the letters or contours filling the regions, then the same effect should occur when the binocular forms are dissimilar. It is very difficult to demonstrate such an effect. Patterns similar to that shown in Fig. 2 were prepared with the binocular inner rectangular array comprised of independent sets of letters in the two eyes. The depth-effect could not be observed, which may have been because the rivalry in the inner region made motor

A	B
	pabcdefgthjklpqw
qwerhjklpmnbva	rasdfghjklpoiuy
poiuxzasdzxcvm	xmnbvcxzasdfghj
plmnyuiopqwerq	mqwertyuioplkjh
zxcvgfdsaqsdfl	onlkjhgfdsaxwer
qwertyuioplkjhg	

FIG. 5. STEREOGRAM WITH SIMILAR INNER BINOCULAR ARRAY, DISSIMILAR OUTER BINOCULAR ARRAY, AND TWO UNIOCULAR COLUMNS AND ROWS

fusion of the different letters impossible. The dissimilar forms could not remain in register because of shifts of convergence, and the effect may, therefore, have been inhibited.

To test for depth with dissimilar superposed forms, the following demonstration was prepared: The inner region of Fig. 5*a* is comprised of a matrix five letters wide and four rows high. The same grouping is reproduced in Fig. 5*b* which is the opposite side of the stereogram. These identical inner matrices furnish the stimulus for motor fusion and permit the patterns to remain in register when viewed in the stereoscope. Surrounding the inner arrays are sets of independently selected letters. There are five columns of these letters to the left of the inner matrix in Fig. 5*b* and four columns to the left of Fig. 5*a*. When superposed, the four

and Wiesel also show, however, that when distinctly different stimuli are imaged in the two eyes, they will be antagonistic. The response of the binocularly fired cortical cell will be inhibited. This process is presumed to be a corollary to phenomenal binocular rivalry. It has been proven that it is possible to obtain stereopsis both in the presence of rivalry and with distinctly different stimuli, *e.g.*, different letter-images. One therefore wonders if it is possible, on the basis of current physiological evidence, to support the contention that stereopsis results from processing a fused field. Nevertheless, this proposal is taken seriously in the text.

columns to the left of 5*a* are in register with four of the columns to the left of 5*b*. This leaves the fifth column to the left of 5*b* to serve as the equivalent to the uniocular column in the previous demonstrations. Similarly, there are four sets of superposed dissimilar columns on the opposite side of the superposed arrays with a fifth column, namely the third column to the right of 5*a*, to serve as the uniocular input.

If it is correct that fusion occurs independently of individual forms, being mediated by the regions which they fill or even by their columnar organizations, then the entire binocular inner pattern should be seen behind the uniocular annulus. If, on the other hand, the letter-forms do play an important role, then only that portion of the inner region containing identical forms in both eyes should go into depth. This is, in fact, what happens, as the reader may judge by viewing the figure in a stereoscope. The dissimilar binocular images are seen in the same plane as the uniocular columns, while the inner similar binocular images are seen in depth. This impression was confirmed by five *O*s. If a fusion-theory must be maintained, then the uniocular image is acted upon by at least one of the inner images which are four columns removed from it. The depth-effect in this demonstration was obtained by the author with as much as 4° of separation between the uniocular image on one side and the nearest column of binocular letters formed from fixation of identical monocular inputs.

Experiment II

To obtain some indication of the stability and magnitude of this phenomenon, the following experiment was conducted.

Method. The method employed was similar to that of the first experiment. The only difference was that *O* manipulated the distance of the red spot himself, controlling it with a lead screw driven by a DC-motor. The light-source was attached to the lead screw. *O* turned a knob to control the rate of displacement of the source until it appeared to be in the same plane as one of the two planes in the pattern. He also determined the direction of motion, toward or away from him, by means of a toggle switch which changed the direction of motor-rotation. The spot always was superposed on the inner region of the stereogram.

The stereogram used is illustrated in Fig. 6. This stereogram was selected because it was found empirically to lead to less frequent fixation-disparity under the viewing conditions than the one shown in Fig. 5. The stereogram also was larger than the one used in Experiment I. The center-to-center separations of the capital letters were about 1.0 cm. The separations of the lower-case letters were about 0.8 cm. The larger size also seemed to reduce the apparent strain involved in maintaining good motor fusion. The viewing distance remained 165 cm.

The procedure was for *O* to fixate the inner letters, move the spot well behind the farthermost plane of the stereogram, and then advance it toward himself until it

was in the same plane as that portion of the display. He then continued to advance the spot until it was aligned with the closest plane in the stereogram. This procedure was reversed when *O* started well in front of the nearest plane and then drove the spot back into alignment with it. After *E* took a reading on the meter-stick, *O* continued to move the spot back until it was aligned with the farthermost

qwerHJKLPmnbva

poiuXZASDzxcvm

plmnYUIOPqwerq

zxcvGFDSAqsdfl

qwertyuioplkjhg

abcdefgthjklpqw

asdfgHJKLPoiuy

mnbvcXZASDfghj

qwertYUIOPlkjh

lkjhpGFDSAlwer

FIG. 6. STEREOGRAM USED IN EXPERIMENT II

plane. The plane which was farthermost depended upon the orientation of the polaroid spectacles. In one orientation the inner binocular letters were behind the mixed binocular letters and the uniocular ring. This relation was reversed for the other orientation of the filters.

In all, 16 measures were made of the alignment with the inner square when it was behind and 16 measures of the alignment with the outer region. When the inner square was forward, another 16 measures were made, and, correspondingly, 16 with the outer region.

Only two *O*s participated in this experiment, the author and another sophisticated *O*. Each *O* served as the other's *E*.

Results. The results of this experiment are consistent with the qualitative observations. One *O* (LK) had a mean distance between the forward and rear planes of the stereogram of 8.8 cm. ($SD = 1.24$ cm.) when the inner region was seen behind. With the inner region forward, the mean depth was 9.7 cm. ($SD = 1.46$ cm.). The second *O* (CP) had a mean depth was 7.5 cm. ($SD = 2.42$ cm.) with the inner region behind, and 10.9 cm. ($SD = 1.47$ cm.) in front of the outer region.

In addition to the foregoing systematic observations, a brief set of measurements of the depth-difference between the mixed binocular letters and the uniocular letters also was made. The spot was so displaced laterally that it would be superposed upon the mixed binocular letters. After several trials in which measures were made of the distance to the mixed letters, and also to the uniocular letters, it became apparent that no depth-difference either could be seen or measured. This failure confirmed the qualitative observation that the mixed letters do not participate in the depth of the identical superposed inner letters.

SOME THEORETICAL IMPLICATIONS

A projection-field model consistent with the classical fusional theory was

published recently by Dodwell and Engel.[17] This model is virtually identical with one published ten years earlier by Linksz.[18] Moreover, two years before the date of Linksz' book, Charnwood published the same model independently.[19] Boring anticipated all these authors by publishing a version of the model in a footnote to *Physical Dimensions of Consciousness*.[20] All of these authors used very similar diagrams to illustrate their model.

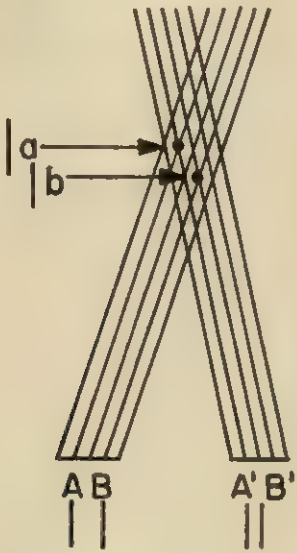

FIG. 7. BASIC STRUCTURE OF PROJECTION-FIELD MODEL
Lines *A* and *A′* can fuse only at *a* and *B* and *B′* at *b*. Places of fusion are at different depths in field.

A version of it is shown in Fig. 7. This model reflects a kind of theoretical consensus concerning stereopsis.

Fig. 7 shows a field or network within which the images arising in the two eyes can be brought into register even when these images fall on disparate places. The geometry is identical with the picture one would draw to show how images in the nuclear plane become arranged into depth when they are projected into the third dimension (Fig. 8). Thus,

[17] C. P. Dodwell and G. R. Engel, A theory of binocular fusion, *Nature*, 198, 1963, 39-74.
[18] Linksz, *op. cit.*, 397-403.
[19] J. R. B. Charnwood, *Essay on Binocular Vision*, 1951, 106-108.
[20] Boring, *The Physical Dimensions of Consciousness*, 1933, 118-119.

doubly represented forms or points in the nuclear plane will become single, or fused, at determinate places in the tri-dimensional space into which they are mapped. The same thing happens in the so-called projection field, which is a neurological representation of binocular space. The model

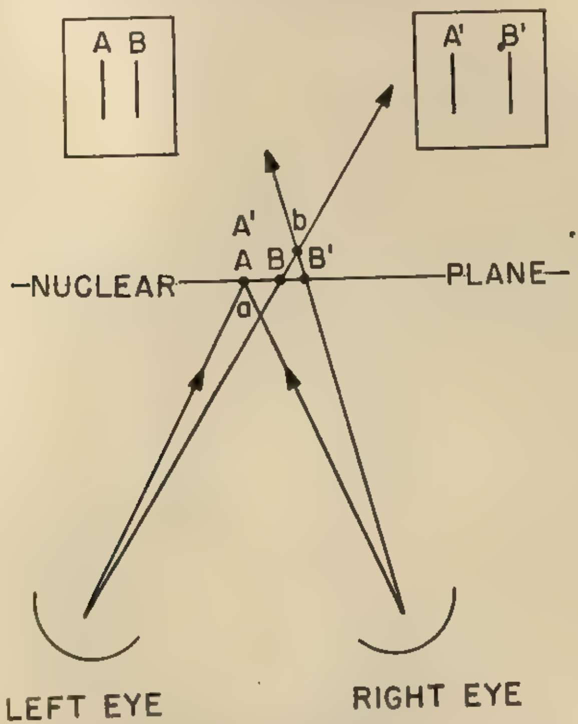

FIG. 8. METHOD OF DRAWING STEREOGRAMS TO ILLUSTRATE PROJECTION THEORY

of the projection-field, then, is a modern-day variation of the old projection-theory of Johannes Kepler.[21] It has simply been flopped over and placed inside *O*'s head.

When Boring presented the model, he stated that it had many limitations. He questioned its physiological possibility and also stated that it was much too simple to be probable.[22] Linksz, however, in describing essentially the same model, suggested that Layer IV of the visual cortex may be the

[21] Boring, *Sensation and Perception in the History of Experimental Psychology*, 1942, 227; see Herman von Helmholtz (*Handbook of Physiological Optics*, Vol. III, J.P.C.S. Southall (ed.), 1963, 483) for a discussion of Kepler's theory and also for an early and charming fragment from Christian Huygen's writings, in which fusion and correspondence are discussed.

[22] Boring, *The Physical Dimensions of Consciousness*, 119.

actual structure fulfilling the functions described above. In addition, he proposed that the stripe of Gennari may serve as the 'fovea' for this cyclopean cortical projection field.[23] Dodwell and Engel go a little further than their predecessors in adding a central recognizer to the projection field. They suggest that, when two retinally mismatched signals fuse in the network, a new signal is transmitted to the central analyzer to be analyzed for pattern and depth.[24]

As already indicated, the model of the projection-field postulates a fusion of signals from the two eyes, but this fusion does not always occur phenomenally. Nevertheless, a fusion of the neural representations of the input is demanded. What is it, then, that is encoded to fuse? This is where the projection-field model breaks down. In most writings on the subject, the authors are satisfied to specify abstract object-points as the basis for signals which fuse. Implicit in much of the earlier work on the subject, however, is the notion that the abstract object-points must be on similar forms or contours.[25] The same conclusion was drawn by Koffka when he asked this same question.[26] We have seen here, however, that it is not necessary for the disparate images to have similar forms. Hence, we cannot say that form-similarity is absolutely necessary. Moreover, it was previously shown that disparity between groupings of dissimilar forms will yield a depth-effect, as will brightness-disparity carried by similar but non-disparate forms.[27] Julesz, too, rejected form as a necessary condition for stereopsis.[28] It seems clear that if a projection-field is involved in stereopsis, it must be associated with an enormously complicated data-processing system which detects common properties of the two half-fields. These properties must exist in both half-fields and must fuse at unique places in the projection-field for depth to occur. Clearly, in this account

---

[23] Linksz, *op. cit.*, 397-403.

[24] Dodwell and Engel, *op. cit.*, 40.

[25] Ogle (*op. cit.*, 274) states that the stereoscopic effect depends upon stimulation from spatial contours.

[26] Koffka, *op. cit.*, 271-272.

[27] Kaufman, *op. cit.*, 398-399.

[28] Julesz (*op. cit.*, *Science*, 145, 1964, 360) claims that contour is not necessary for stereopsis, but his demonstrations are not critical in this regard. In none of his demonstrations has he eliminated similar contours surrounding dot-clusters in the two half-fields. It is well-known that with various transforms, as with meridional lenses and aneisokonia, it is possible to experience stereopsis. Although the contours are distorted, they still are present, and stereopsis results. The same holds for Julesz's patterns. His claim may, instead, be founded upon the notion that the over-all form seen in depth is not present in either half-field, but this is not surprising since large numbers of proximate clusters with contours do go into depth. These clusters are bound to organize as an inner rectangle. (C. S. Bridgman, Analysis of 'A recently discovered stereoscopic effect,' this JOURNAL, 77, 1964, 138-143.)

of the matter, the uniocular image will be as likely to fuse with one image from the central region of Fig. 5 as any other. It can, therefore, have no unique reference-position, or place of fusion, in the projection field, and should, therefore, be indeterminately located in space; but it is not. It is seen either in front of, or behind, the inner pattern of like stimuli, and in the same plane as the unlike binocular stimuli. Simple fusion of image-representations in a network, cannot, then, be the basis for depth-perception.

As an alternative to the classical view, Julesz suggested that the system may be responding to a binocular field made up of representations of differences in brightness of micropatterns or point-clusters in the monocular fields.[29] The difference-field will contain a distinctive region corresponding to the binocular form which is seen in depth. Two such representations of the form must be present; one form when the monocular inputs are in register, and another form when they are shifted or convolved relative to each other in the amount and direction of the binocular parallax. In Julesz's regular dot-patterns, an inner square is seen in the binocular difference-field when the monocular patterns are in register, and also when they are shifted by as much as the disparity between the two patterns. It is obvious that the difference-fields correspond to fused images, and the shifts get them to correspond to different depths in the projection-field model. In the present demonstrations, there are not two unique difference-fields to define the inner form and its depth. A single difference-field, when the patterns are in register, does contain the inner form. Physically convolving the monocular patterns does not, however, produce a second difference-field with a unique form corresponding to the binocular form seen in depth. In Julesz's model this is necessary for the depth-effect. Yet Julesz's model seems to be in the right direction. He recognizes that it is but a crude approximation to a more suitable model. He also acknowledges that brightness-subtraction is a highly questionable operation for inclusion in the ultimate model.

One may agree with Julesz that there is, in the most general sense, a requirement for some form of correlation of the monocular inputs, but the operations performed on the correlated inputs must be essentially logical in nature. Because one kind of operation yields an index of correlation between inputs, we cannot conclude that other correlations do not

[29] Julesz, Towards the automation of binocular depth perception, in *Proceedings of IFIP Congress 62, Information Processing 1962*, 1962, 439-444; Julesz, Binocular depth perception of computer-generated patterns, *Bell System Tech. J.* 39, 1960, 1151-1156.

exist. It is unlikely that all relevant associations between monocular stimuli can be revealed by linear algebraic operations such as subtraction or addition. The essential relations between stimuli must be common organizations and their positions in perceptible spatial patterns. This is essentially a matter of detecting a difference in the phases of the correlated stimuli with respect to a spatial reference-system. Let us see how this general view applies to the demonstrations in this paper.

In Demonstration I, the inner identical 'fused' letters went into depth relative to the uniocular letters. There are two places of agreement or correlation in the half-fields. There is, first, the identity of the superposed letters. Secondly, there is the identity of the sizes of the over-all patterns. Since the correlated letters occupy different places in their respective frameworks, they are not in phase. This phase-difference is not a disparity, because it is not dependent upon differences in retinal location. Instead, it is dependent upon position in a frame of reference, *i.e.* the over-all pattern.

In Demonstration II, the correlated letters differ in position with respect to one edge of the over-all patterns, and do not differ with respect to the opposite edge. There is, therefore, a phase-shift with respect to but one edge. This, too, gives a depth-effect, which suggests that edges of patterns, as well as over-all sizes of patterns, and correlations of pattern-elements can interact to produce the stimulus to depth.

In Demonstration III, the inner correlated region goes into depth with respect to the relatively uncorrelated rivalrous columns of letters. Moreover, the uncorrelated letters appear to be in the same plane as the uniocular letters. This result was predicted in advance on the basis of the fact that the correlated letters occupied different places in their respective over-all sequences. This difference in place constitutes the phase-difference. It may be of some interest to note that phase, in this sense, is similar to the notion recently set forth by Rock and Ebenholtz in their discussion of the *phi*-phenomenon, and by Rock and Brosgole in their work on proximity as an organizing factor.[30] In these cases it was found that positions of the stimuli as perceived, instead of their positions on the retina, can determine both stroboscopic motion and organization. There also is a similarity to Wallach and Lindauer's position that form-properties are depth stimuli.[31] The important point here is that relative position in an over-all pattern

[30] Irvin Rock and Sheldon Ebenholtz, Stroboscopic movement based on change of phenomenal rather than retinal location, this JOURNAL, 75, 1962, 193-207; Irvin Rock and Leonard Brosgole, Grouping based on phenomenal proximity, *J. exp. Psychol.*, 67, 1964, 531-538.

[31] Hans Wallach and Judith Lindauer, On the definition of binocular disparity, *Psychol. Beitr.*, 6, 1962, 521-530.

can determine stereopsis. This is a more general notion than disparity, which is a retinal concept, but it includes disparity. All similar disparate stimuli differ in phase, but not all stimuli which differ in phase necessarily exhibit disparity, at least in the sense of normal binocular vision where the disparate objects viewed are the same for both eyes. This is true in all the demonstrations described in this paper.

Demonstration IV

One way in which to test the generality of the phase-difference concept is to see how it applies in ordinary stereograms. The stereogram in Fig. 9 is a case in point. There is an inner region of identical letters in both members of the stereo pair, and the background letters of both members are the same. Consequently, there is no disparity between any of the inner letters relative to most of the background letters. Nevertheless, the inner region of the right-hand member of the pair is displaced one letter-space to the left of the over-all pattern, which serves to create a disparity between the inner regions relative to the edges of the patterns.

If the inner regions of *E*s and *W*s were presented alone, *i.e.* without the background *A*s, and within line-rectangles of equal size, then the same sort of disparity would produce an immediate and compelling depth-effect. The inner array of letters would be seen behind the surrounding rectangle. In this case, however, the depth-effect will only occur when the edge of the binocularly viewed pattern is fixated and brought into register. As a matter of fact, with a little practice it is possible to shift fixation from the edge of the pattern to the inner region and, with an appropriate convergence-change, see the inner region go into and out of depth. As the inner region achieves good registration, it becomes coplanar with the *A*s. When it is out of register, with the edges in register, it moves behind the background. This impression was confirmed by three *O*s in the following way: The stereogram was placed in the polaroid projector and so adjusted that the inner region was out of register on the ground-glass screen, while the edges of the pattern were in register. The *O*s reported depth soon after looking at the pattern. Then *E* turned a knob which affected the registration of the two half-fields. The inner regions were brought into register. This procedure was repeated several times. All *O*s reported a loss of depth when the inner region was brought into register, and the appearance of depth when the relation was reversed.

This example illustrates the phase-difference notion. When the inner region is in register, then, it is in phrase with equally well correlated background-letters. It should, therefore, be seen in the same plane as all of the

background-letters excepting those at the edges of the pattern. Here the single vertical columns are in front of the over-all array as in the earlier figures. When, however, the inner letters are out of register, they alone will go into depth relative to the background-letters. The background-letters are perfectly correlated spatial sequences with periods of unity. Hence any change of registration of these letters by the amount of one letter-space will bring them back into phase. This is not true of the inner array. A shift of one monocular inner grouping by one letter-space relative to the other will create a phase-difference relative to the adjacent background-letters, which is due to the fact that the inner region does not have a sequence-length of one letter-space. Hence the depth-effect will occur, and its occurrence will depend upon the registration of the two half-fields.

Demonstration V

Fig. 10 shows a stereogram similar to that in Fig. 9. The only difference is that the background-letters all are dissimilar and, therefore, not well

```
AAAAAAAAAAAAAAA          AAAAAAAAAAAAAAA
AAAAAAAAAAAAAAA          AAAAAAAAAAAAAAA
AAAAAEEWWEAAAAA          AAAAAAEEWWEAAAA
AAAAAEWEWWAAAAA          AAAAAAEWEWWAAAA
AAAAAWEEWWAAAAA          AAAAAAWEEWWAAAA
AAAAAEEWWEAAAAA          AAAAAAEEWWEAAAA
AAAAAAAAAAAAAAA          AAAAAAAAAAAAAAA
AAAAAAAAAAAAAAA          AAAAAAAAAAAAAAA
```

Fig. 9. Stereogram with Identical Elements and Inner Region Shifted with Respect to Edge

```
TTTTTTTTTTTTTTT          AAAAAAAAAAAAAAA
TTTTTTTTTTTTTTT          AAAAAAAAAAAAAAA
TTTTTEEWWETTTTT          AAAAAAEEWWEAAAA
TTTTTEWEWWTTTTT          AAAAAAEWEWWAAAA
TTTTTWWEEWTTTTT          AAAAAAWWEEWAAAA
TTTTTEWWEETTTTT          AAAAAAEWWEEAAAA
TTTTTTTTTTTTTTT          AAAAAAAAAAAAAAA
TTTTTTTTTTTTTTT          AAAAAAAAAAAAAAA
```

Fig. 10. Stereogram with Different Background Elements and with Inner Region Shifted with Respect to Edge

correlated with regard to form or contour. The inner letters all are alike but, again, one set is closer to one edge of the over-all pattern than the other. In this case, the inner region is seen to be behind the background-letters. Moreover, the depth-effect will occur when the inner region is in

register, even though that does not happen with any reliability in the previous stereogram.

The difference between the two stereograms, then, is that the inner letters will not go into depth when they are in register in Fig. 9 but will go into depth when they are in register in Fig. 10. Presumably, the fact that the background-letters are well correlated in one pattern and not in the other underlies this difference.

The phase-difference theory can account for this difference in result. In Fig. 9, there is no phase-difference between the background-letters and the inner letters when the inner letters are in register. Hence the background-letters are seen as coplanar with the inner letters. In Fig. 10, however, the background-letters are poorly correlated with each other regardless of their registration. It is quite pointless to talk about the phase-relations between uncorrelated sequences. When the inner letters are in register, they are out of phase with respect to the over-all outlines of the patterns. This phase-difference will mediate the depth-effect, then, even when the inner patterns are in register, which cannot happen in Fig. 9 because the background-letters are redundant, and changing fixation will not change their phases.

Conclusions

The demonstrations described in this paper suggest two major conclusions. The first is that sensory fusion of disparate stimuli is an inadequate concept for the prediction of stereopsis. The second conclusion is that the concept of disparity requires a more careful definition.

Disparity exists between two points when they occupy different visual directions. Disparity is an adequate stimulus for stereopsis when there are other points with different disparities in the visual field. This concept of disparity breaks down, however, when the points are contained within relatively complex and different visual fields. One has to ask the question, "Disparity of what?" In Demonstration I, the disparity is between one of the binocular images and uniocular image, but the same disparity does not produce a depth-effect in Demonstration III when the superposed binocular images are dissimilar. This suggests that disparity between arbitrary points or objects is not the basis for stereopsis. The nature of the disparate objects must be included in the definition of the stimulus to depth.

The concept of fusion is itself a means for including the nature of the object in a definition of effective disparity; Only similar objects can fuse and, therefore, only similar objects can be seen in depth when they are dis-

parate. We have seen, however, that this usage of 'similar' fails. The whole fusion-notion becomes untenable in the light of Demonstration III. Fusion should occur across all of the binocular forms, but it does not. Depth only occurs for the similar forms, and not for the dissimilar forms. The fusion-concept also fails to predict the outcome of Demonstration V, where depth occurs without any disparity except on the outer edge of the patterns, which is sufficient to cause a correlated inner region to go into depth with respect to background. Fusional models do not predict that only the correlated inner array will be affected by the edge-disparity. The background will partake of the depth depending upon how well correlated the two half-field backgrounds are.

All of this suggests a somewhat different definition of the stimulus to stereopsis. Stereopsis will occur when correlated stimuli are out of phase with respect to some reference-system. The reference-system can be another set of correlated objects or point-arrays, or it can be the edges of the over-all half-fields. Any mechanism which can detect the correlations between the binocular stimuli and also detect a difference in their phases can yield of representation of depth. Fusion is not necessary.

Several major problems remain to be solved. One is the problem of defining the stimulus-dimensions which, when correlated, lead to depth-perception. Another is the problem of specifying the operations by which the system determines the existence of a correlation. It is suggested that binocular rivalry and, particularly, the spread of suppression, may play a role in mediating the matching process and in detecting phase-differences.

INTEROCULAR TRANSFER OF PATTERN-DISCRIMINATION IN THE GOLDFISH

By Sandra Milstein Shapiro, Bryn Mawr College

Recent research on interhemispheric relations, of which there are several reviews,[1] has played a promising role in the advancement of our knowledge of brain-function, particularly with respect to the problem of how information stored in one part of the brain becomes available to another part. To find out how information normally is stored and transmitted, connections between sources of input and various areas of the brain are systematically eliminated until transfer is affected. In the visual modality, with which the present paper is concerned, interocular transfer has usually been studied by training with one eye and testing with the other. Restricted distribution of the original sensory input, or of subsequent elaborations and recodings of this input, may be accomplished (1) by occluding one eye; (2) by cutting those optic fibers which, in mammals, cross at the optic chiasma, or by choice of inframammalian organisms whose optic fibers project unilaterally; (3) by systematically destroying the fiber-tracts which cross the midline at various levels, or by choice of organisms which lack some of these crossing tracts.

The results of systematic midline-surgery in mammals suggest that the optic chiasma corpus callosum, and possibly other forebrain-commissures are necessary for interocular transfer of pattern-discrimination,[2] although there is one report of successful transfer in split-brain cats when shock-avoidance is substituted for the usual food-reinforcement.[3] Easier brightness-discriminations appear to survive sec-

* Received for publication May 30, 1964. This paper is adapted from a doctoral dissertation submitted to Bryn Mawr College. The author is indebted to Professor M. E. Bitterman, under whose direction the work was done. The costs of the research were met by Grant MH-02857 from the Public Health Service.

[1] R. W. Sperry, Cerebral organization and behavior, *Science*, 133, 1961, 1749-1757; V. B. Mountcastle (ed.), *Interhemispheric Relations and Cerebral Dominance*, 1962, 1-294.

[2] R. E. Myers, Transmission of visual information within and between the hemispheres: A behavioral study, in Mountcastle (ed.), *op. cit.*, 51-73; J. L. de C. Downer, Interhemispheric integration in the visual system, in Mountcastle (ed.), *op. cit.*, 87-100; C. R. Hamilton and M. S. Gazzaniga, Lateralization of learning to colour and brightness discriminations following brain bisection, *Nature*, 201, 1964, 220.

[3] J. A. Sechzer, Successful interocular transfer of pattern discrimination in "split-brain" cats with shock avoidance motivation, *J. comp. physiol. Psychol.*, 57, 1964, 76-83. (The above discussion of Sechzer's experiment is based on the paper as read at the meeting of the Eastern Psychological Association, 1962.)

tion of the optic chiasma and corpus callosum in cats, but more difficult discriminations fail to transfer.[4] Failure of transfer of color- and brightness-discriminations in split-brain monkeys is reported in several studies in which a relearning measure is used,[5] but there is some indication of transfer when conflicting color- or brightness-discriminations are presented simultaneously to the two eyes.[6]

In the visual systems of the submammalian vertebrates, there is complete decussation at the optic chiasma with fibers from one eye terminating exclusively in the contralateral optic tectum. The neocortex and the accompanying corpus callosum are absent. While those commissures which appear to function in the transfer of pattern-discrimination in mammals are not present, there are other commissures in the forebrain, 'tweenbrain, and midbrain which might serve an analogous function. For students of comparative psychology interested in phyletic changes in the mechanisms of memory, discrimination, generalization, and transfer, it is important to establish the extent of interocular transfer in these species. If transfer is present among submammalians, then it is clear that the brain has undergone functional reorganization in this respect. The optic chiasma corpus callosum, and anterior commissure may be implicated in the establishment of binocular pattern-equivalence in the mammal, but the submammalian utilizes other commissures and areas of the brain for this purpose.

Working with pigeons, Levine observed interocular transfer of color-, brightness-, and pattern-discrimination when stimuli were presented anteriorly, within the area of binocular overlap. Transfer was absent when stimuli were presented to the lower temporal retina, which never participates in binocular vision. Levine also reported a preliminary experiment in which cutting the commissure between the superior colliculi resulted in a failure of interocular transfer under conditions which would have given transfer in normal pigeons.[7]

Catania suggested that Levine's negative results with lateral stimuli were an artifact of his procedure.[8] Levine used a modified Lashley jumping stand in which, according to Catania, the pigeons would have had to cock their heads to the side when jumping. With the original eye covered in the transfer-tests, the habit of cocking the head in the now incorrect direction interfered with performance until a new bodily orientation was learned, making it appear as though transfer had failed. In Catania's study, pigeons with goggles permitting either anterior or lateral vision were trained monocularly on color-, brightness-, and pattern-discriminations. The *Ss* were reinforced for key-pecking in the presence of the positive stimulus on a *VI*-3 min. schedule and unreinforced in the presence of the negative stimulus.

---

[4] T. H. Meikle and J. A. Sechzer, Interocular transfer of brightness discrimination in "split-brain" cats, *Science,* 132, 1960, 734-735; Meikle, Role of corpus callosum in transfer of visual discriminations in the cat, *idem,* 1496.

[5] Downer, *op. cit.;* Hamilton and Gazzaniga, *op. cit.,* 220.

[6] C. B. Trevarthen, Double visual learning in split-brain monkeys, *Science,* 134, 1961, 258-259.

[7] J. Levine, Studies in the interrelations of central nervous structures in binocular vision: I. The lack of bilateral transfer of visual discriminative habits acquired monocularly by the pigeon, *J. genet. Psychol.,* 67, 1945, 105-129; II. The conditions under which interocular transfer of discriminative habits takes place in the pigeon, *J. genet. Psychol.,* 67, 1945, 131-142.

[8] A. C. Catania, Techniques for the control of monocular and binocular viewing in the pigeon, *J. exp. Anal. Behav.,* 6, 1963, 627-629.

Transfer-tests were carried out during extinction. When the anterior goggles were worn, interocular transfer was obtained for brightness and color, but not for pattern. Catania noted that pigeons are near-sighted for anterior stimuli and far-sighted for lateral stimuli, and that his lateral patterns, projected only 1-2 in. from the eye, probably were out of focus.

In birds, then, there is evidence for the interocular transfer of brightness-, color-, and pattern-discriminations between the retinal areas of binocular overlap, and transfer at least of color- and brightness-discriminations between the nonoverlapping areas of the retina. Catania's study points to the importance of controlling peripheral factors in the assessment of transfer.

Interocular transfer in fish first was studied by Sperry and Clark.[9] Sixteen Goby fish were trained monocularly on an object-discrimination, with visual and spatial cues confounded. Transfer-tests following overtraining revealed considerable inter-individual variability. About half the *Ss* showed little or no transfer, three *Ss* showed an increase of only one error in the 25-trial transfer-sessions, and the remaining *Ss* performed at an intermediate level. While Sperry and Clark concluded that the neural mechanism for transfer is not perfectly developed in fish, the variability of their results might have resulted also from damage to the transfer-eye, which was covered during the relatively long period of original training; to the confounding of visual and spatial cues, which may transfer in different degrees; or to orientational and motor habits learned during original training which would be appropriate for locomotion and aiming when one eye was used but inappropriate when the other eye was used.

Schulte reported significant interocular transfer of pattern-discriminations in 18 of 21 carp trained to a high level (90-99% correct choice during the last 100 training trials) prior to transfer-testing.[10] The *Ss* previously had undergone binocular exposure to the testing stimuli in preference-tests, a procedure which may have facilitated transfer.

From the results of a series of experiments on interocular transfer in the goldfish, McCleary concluded that discrimination of patterns does transfer, and that negative results are due to the failure of visuomotor learning to transfer.[11] When the response-measure was cardiac deceleration, which presumably involved no skilled response, interocular transfer was observed in the form of differential cardiac deceleration in the presence of shocked as compared with neutral stimuli. When, however, a new instrumental response was required, transfer was found only when the naïve eye, although not exposed to the discriminative stimuli, was uncovered. According to McCleary, the naïve eye must learn the response-requirements for *S* to be able to perform in transfer-tests, since visuomotor learning does not transfer.

From the studies of interocular transfer in the fish, it may be concluded that a significant degree of transfer does occur in some learning situations. The completeness of transfer has not, however, been assessed in a study in which there were

---

[9] R. W. Sperry and Eugenie Clark, Interocular transfer of visual discrimination habits in a teleost fish, *Psychol. Zoöl.*, 22, 1949, 372-378.

[10] A. Schulte, Transfer und Transpositionversuche mit monokular Dressierten Fischen, *Z. vergl. Physiol.*, 39, 1957, 432-476.

[11] R. A. McCleary, Type of response as a factor in interocular transfer in the fish, *J. comp. physiol. Psychol.*, 53, 1960, 311-321.

controls for general learning (*e.g.* adaptation to the experimental situation). In cases in which failure of transfer has been reported, there is some suggestion that response-factors may have been responsible. If motor difficulties could be eliminated by monocular response-training with each eye, then a purer measure of the amount of sensory transfer might be obtained for any visual discrimination subsequently introduced.

Experiment I

The main purpose of Experiment I with goldfish was to obtain some objective measures of the level of interocular transfer—complete, partial, or absent—of a visual pattern-discrimination, with appropriate controls for transfer of general learning and adaptation to the experimental situation, and for the response-biases which may accompany monocular vision. Two features of the design ensured equivalent experience with each eye in the non-discriminative aspects of learning. First, the *Ss* were trained on a binocular color-discrimination and then given sessions of monocular target-practice with each eye prior to the interocular-discrimination stage of the experiment. Secondly, since training days with the right eye and left eye were interspersed during the experiment proper, the general condition and state of adaptation of each animal was equal for right- and left-eye sessions. A failure of transfer under these conditions could be interpreted only as a failure of discriminative transfer and not as a failure of response-transfer or as a result of a change in the state of the organism.

The plan to distinguish complete from partial transfer, and partial transfer from zero transfer was based upon the following rationale: If there is no interocular transfer at all, then learning should proceed independently with each eye, and it should not matter whether the animal learns the *same* discrimination with each eye (Nonreversal) or *opposite* discriminations with two eyes (Correlated-Reversal). To the extent that Nonreversal (*NR*) performance is superior to Correlated-Reversal (*CR*) performance, there is some interocular transfer or lack of independence in learning with each eye. To evaluate the *completeness* of interocular transfer, *CR* performance is compared with Uncorrelated-Reversal (*UR*) performance. In the *UR* condition, the animals encounter the same number of discrimination-reversals as in the *CR* condition, but the relation between the positive stimulus and the eye used is 'random.' If interocular transfer is less than complete, *CR* performance should be superior to *UR* performance.

Relative performance under the three conditions as predicted on the assumption of interocular transfer at each of the three levels may be

summarized as follows: (1) If there is *no* interocular transfer, $NR = CR > UR$; (2) If there is *partial* interocular transfer, $NR > CR > UR$; (3) If there is *complete* interocular transfer, $NR > CR = UR$.

Subjects. The *Ss* were 23 goldfish, about 3 in. in length from the head to the base of the tail, obtained from a local dealer.

Blinders. The blinders used to occlude vision were made of opaque black Vinylite (manufactured by the Transilwrap Company of Philadelphia), 0.02-0.03 in. thick before molding. An adequate range of sizes was produced by varying the diameter of the punch-and-die used to mold the blinder in a drill press. A small piece of Plexiglas was glued to the outside of the blinder to provide a handle which facilitated insertion and removal.[12] The entire blinder was then painted wtih a double coating of airplane dope (Testor's Butyrate Dope, gloss black No. 19) to insure opacity after the material had been stretched during the molding process, and the edges were finished by filing them with a fine sandpaper and polishing with crocus cloth. The blinders were thoroughly washed before use.

The initial fitting of the blinder was accomplished after the fish had been lightly anesthetized in a solution of MS222 Sandoz (Tricaine Methanesulphonate) to reduce the trauma of the fitting procedure. A fish wearing a blinder is shown in Fig. 1.

A test of the opacity of the blinders was made for each *S* while it was wearing blinders on both eyes. Following a 10-min. period of dark-adaptation in the experimental chamber, 20 trials were given on the previously learned green-red discrimination. No *S* hit the target during its opacity test. Most *Ss* remained motionless on the bottom of the tank during the entire session, although some *Ss* occasionally were observed to move along the wall of the tank for short distances. When the blinders were removed, *Ss* were observed to strike the appropriate target promptly.

Apparatus. The apparatus consisted of a black Plexiglas enclosure which was equipped with a hinged front and a hinged lid and housed a 2-gal. aquarium into which the fish was placed. When the front and lid were closed, the interior of the enclosure was completely darkened. A circular target of clear Plexiglas, fixed to the lid on a thin rod, was lowered with the lid into a position close to the front wall of the aquarium and about 2 in. below the surface of the water. An In-Line projector was mounted on the enclosure behind the target in such a way that colored lights or patterns could be projected on it. Fig. 2 shows the choice-apparatus used in Experiment II, which was equipped with two targets and two projectors. The apparatus used in Experiment I was the same except that it was equipped with but a single target and projector centered on the left wall.

The response of *S* was to strike at the target. The rod in which the target was mounted was set into the needle-holder of a crystal phonograph cartridge. Contact with the target generated a voltage across the cartridge which, when amplified and integrated, served to operate relays which determined the consequences of the response. The technique is described in greater detail elsewhere.[13] The *Ss* were reinforced with live *Tubifex* worms delivered by an automatic feeder located at the

[12] This feature of the blinder was suggested by Dr. Jerald J. Bernstein.

[13] Nicholas Longo and M. E. Bitterman, Improved apparatus for the study of learning in the fish, this JOURNAL, 72, 1959, 616-620.

end of the tank opposite the target.[14] With each operation of the feeder, the translucent inner rear wall of the enclosure was diffusely illuminated (for 6 sec. during binocular training and for 15 sec. during monocular training) by a small white lamp set behind it, permitting *S* to find the worm.

The programming of the experiment was semi-automatic. *E* manipulated a switch to determine which stimulus would appear on each trial, read the latency of re-

FIG. 1. FISH WEARING A BLINDER IN THE ACT OF RETRIEVING A WORM WHICH HAS BEEN DISCHARGED INTO THE TANK

sponse to the nearest 0.5 sec. from a timer, and reset the latency-timer to zero at the end of a trial. All other features of the technique were automatic.

Adaptation and preliminary training. The *Ss* were brought into the temperature-controlled experimental room and housed there in individual 2-gal. aquaria set on open racks. Each home tank was cleaned and aerated by a bottom filter.

During adaptation, the *Ss* received daily feedings of 15-20 mg. of trout chow.

[14] Longo and Bitterman, An improved live-worm dispenser, *J. exp. Anal. Behav.*, *6*, 1963, 279-280.

They also were fed *Tubifex* worms delivered by an eye-dropper into the home-tank. When they were taking the worms readily in the home-tank, feeder-training was begun. Each *S* was transported to the experimental enclosure in a shallow Pyrex dish filled with water, with *E*'s hand loosely cupping *S* during the transfer, and placed in the experimental tank. The room was darkened, and the front door of the enclosure was left open that *E* might observe *S*'s behavior. The feeder was operated at irregular intervals and the feeder-light, under manual control for this purpose, was left on until *S* had retrieved the worm, or until 15 sec. had elapsed.

When the worms were being taken regularly, target-training was begun. First a

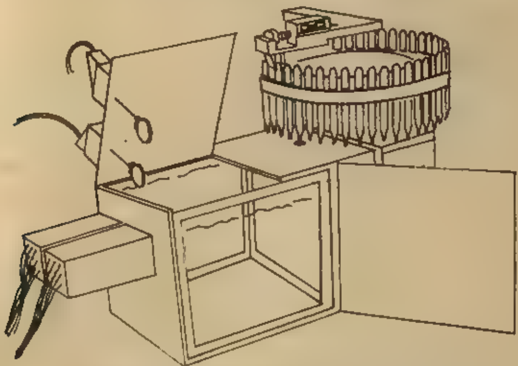

FIG. 2. SKETCH OF THE APPARATUS

(The apparatus shown was used in Experiment II. The apparatus used in Experiment I was the same except that it had only one target and projector, centered on the left wall.)

wire-mesh target baited with a dried mixture of trout chow and water was used. Each trial began with the illumination of the mesh target by a yellow light. Each contact of the animal with the illuminated target turned off the light and initiated a reinforcement-cycle (turned on the feeder-light and delivered a *Tubifex* worm), after which there was a 5-sec. intertrial interval in the darkness. The time during which the feeder-light remained on was decreased progressively until *S* was taking the worm within 6 sec., the time to be allowed during all subsequent binocular pre-training sessions.

When the response to the baited target was well established, an unbaited clear Plexiglas target was substituted for it. On each trial the target (illuminated with yellow light) was presented until *S* hit it, or until 30 sec. had elapsed. As many trials were given each day as were required for *S* to earn 10 reinforcements. After *S* had hit the target on 10 successive trials, which required about six days on the average, an additional 50 trials were given, with which the binocular target-practice was concluded.

In the next phase of preliminary training, the animals were taught a green-red discrimination. Each trial began with the illumination of the target which was

green on half the trials and red on the remainder of the trials, the order of presentation of the two colors following selected Gellermann orders.[35] Some of the *Ss* were trained with red-positive, some with green-positive. On positive trials, response to the target turned off the target-light and initiated the reinforcement-cycle; if there was no response in 30 sec., the target-light was turned off and an intertrial interval followed. On negative trials, the target-light remained on for 30 sec. independently of the animal's response, after which it was turned off, and the intertrial interval began. During this stage of pretraining, the animals were given 20 trials per day for 9 days, by which time the median latency of response to the positive stimulus had stabilized at 15 sec. or less, and the median latency of response to the negative stimulus was consistently 30 sec., the value assigned when no response was made.

The final stage of preliminary training consisted of monocular target-training, the eye over which the blinder was placed being alternated daily. The procedure was the same as in the binocular target-training stage except that the yellow target was used and all responses to it were reinforced. Early in this monocular training it was necessary to increase the amount of time during which the feeder-light stayed on to permit *S* to find the worm. Training was continued until the target-response was reëstablished and *S* was taking the worm within 15 sec., the time to be allotted during the experiment proper.

On the basis of their performance during the last two stages of preliminary training, the *Ss* were assigned to the various experimental and control conditions as set forth below. Of the original 23 *Ss*, six were lost because of illness, and a seventh, which failed to hit the target during monocular target-training, was dropped from the experiment.

Interocular discrimination. The experiment proper consisted of 32 daily sessions of 40 trials each, with a 30-sec. time-limit. Training was monocular; the eye upon which the blinder was placed was alternated daily for some of the animals, while for others it was varied randomly with the restrictions (1) that the number of days on which the left and right eyes were used be equal in each eight-day block, and (2) that the same eye not be covered on more than two successive days. One stimulus was a white horizontal stripe against a dark background, the other was a white vertical stripe against a dark background. The stimuli, as projected on the target, were each ⅛ in. wide × ¾ in. long. The order of presentation of the two stimuli followed the rules of Gellermann.

Of four *Ss* in the *NR* Group, two were trained with the vertical stripe always positive and the horizontal stripe always negative, while the other two were trained with the horizontal positive and the vertical negative. For two of the animals, right and left eyes were alternated daily, while for the other two the more irregular order was used. Of five *CR* animals, three were trained with vertical-positive for the right eye and horizontal-positive for the left, while for the others the relation between eye and positive stimulus was reversed; for three of the five, there was daily blinder-alternation, for the other two irregular blinder-alternation. Of five *UR* animals, three had blinders alternated daily and two irregularly; for each of these animals, of course, the reversal-schedule was such that there were as many days with vertical-positive for the right eye as there were days with vertical-positive for the left eye, and as

[35] L. W. Gellermann, Chance orders of alternating stimuli in visual discrimination experiments, *J. genet. Psychol.*, 42, 1933, 206-208.

many days with horizontal-positive for the left eye as with horizontal-positive for the right eye.

Two binocular control *S*s were also trained, on half the days with vertical-positive and on half the days with horizontal-positive; for one of the *S*s, the reversal was daily, while for the other the order of tasks was the same as in the *CR* condition.

Results. Individual curves showing the median latency of response to the positive stimulus and to the negative stimulus, both for the right eye and for the left eye, are presented in Fig. 3. Over-all differences in learning

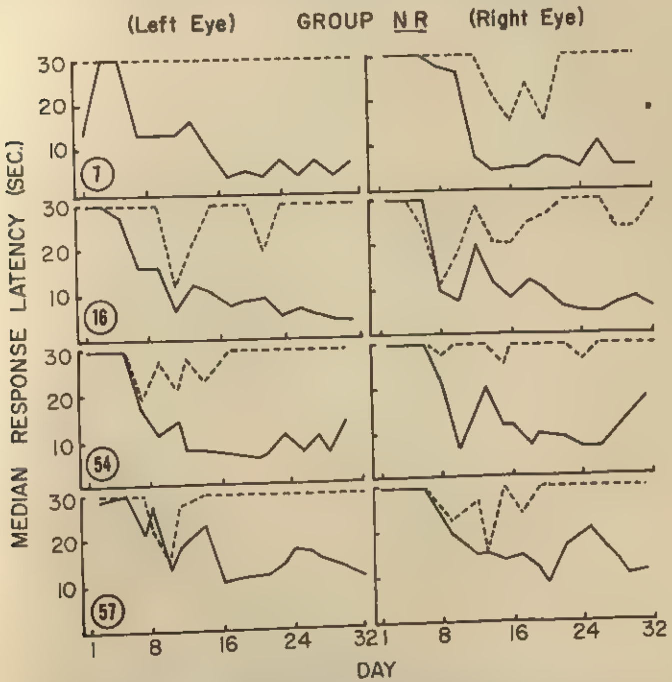

FIG. 3A. INDIVIDUAL CURVES FOR THE *NR* ANIMALS OF EXPERIMENT I SHOWING THE MEDIAN LATENCY OF RESPONSE TO THE POSITIVE STIMULUS (SOLID LINE) AND TO THE NEGATIVE STIMULUS (BROKEN LINE)

among the three groups are summarized in Table I, which shows for each *S* the number of days in each eight-day block on which latencies to the negative and positive stimulus differed significantly. Significance was determined by Wilcoxon's Test for Unpaired Replicates performed on the daily data for each *S*.[16] The data show that each of the *NR* animals

[16] Frank Wilcoxon, *Some Rapid Approximate Statistical Procedures*, American Cyanamid Company, 1949, 1-16.

mastered the discrimination; each animal responded to the two stimuli with significant differential latency on each of the eight days of each of the last two blocks. While there appears to have been some improvement in the scores of *CR* and *UR* animals, particularly in Block 4, the improvement

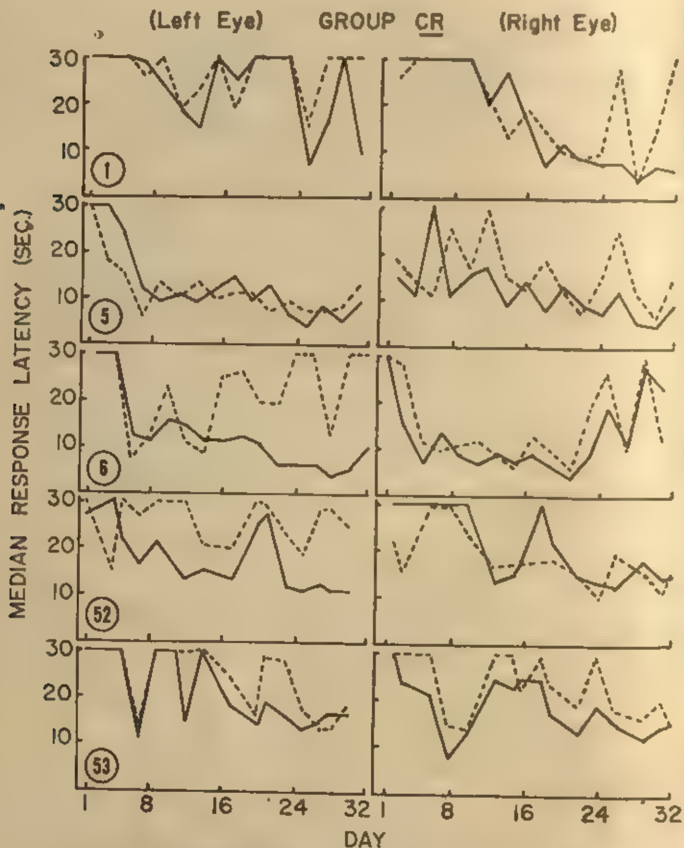

FIG. 3B. INDIVIDUAL CURVES FOR THE *CR* ANIMALS OF EXPERIMENT I SHOWING THE MEDIAN LATENCY OF RESPONSE TO THE POSITIVE STIMULUS (SOLID LINE) AND TO THE NEGATIVE STIMULUS (BROKEN LINE)

was due, in great part, to the development of stimulus-preferences. During the last two blocks of days, significant differences in latency appeared only when *one* of the stimuli was positive and not when the other was positive for *CR* Ss 6 and 52 and for *UR* Ss 14, 15, and 55. The vertical stimulus was positive on five of the six days in which *UR S* 58 showed significant differential latency.

The data in Table I were subjected to an analysis of variance. The two main effects, *Groups* ($F = 68.89$, $df = 2/11$) and *Blocks* ($F = 27.53$, $df = 3/33$), as well as the interaction of *Groups* $\times$ *Blocks* ($F = 7.09$, $df = 6/33$), are significant beyond the 0.01 level. Further tests, by

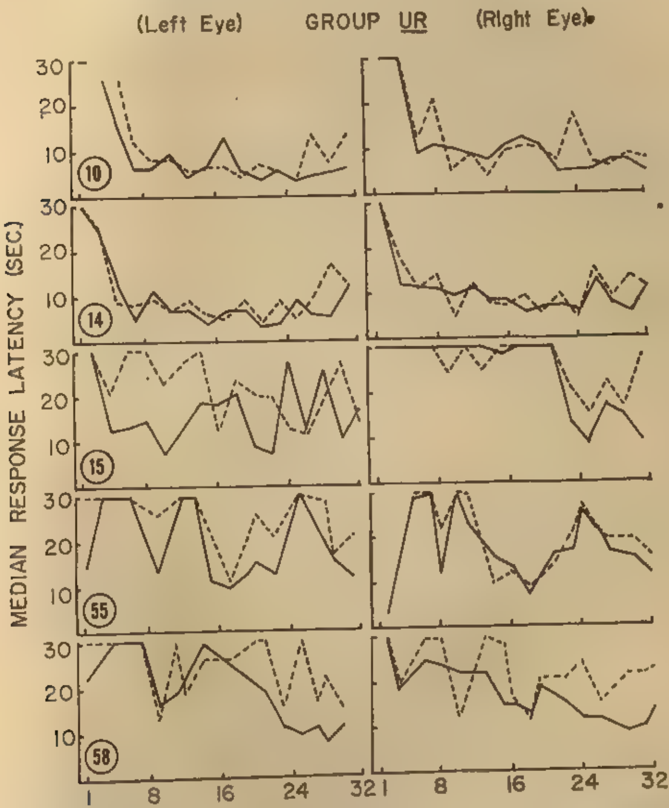

FIG. 3C. INDIVIDUAL CURVES FOR THE *UR* ANIMALS OF EXPERIMENT I SHOWING THE MEDIAN LATENCY OF RESPONSE TO THE POSITIVE STIMULUS (SOLID LINE) AND TO THE NEGATIVE STIMULUS (BROKEN LINE)

Scheffé's method for determining group-differences following a significant overall *F*, revealed that the *NR* group differed significantly ($p < 0.01$) from *CR* and *UR* groups, which did not differ significantly from each other.[17]

While the above analysis supports the hypothesis of complete interocular

[17] Quinn McNemar, *Psychological Statistics*, 1962, 285-286.

transfer ($NR > CR = UR$), it is possible that more subtle differences between the *CR* and *UR* groups might emerge if the occurrence of stimulus-preferences were analyzed. If stronger stimulus-preferences developed in the *CR* group, it would suggest either a partial functional separation of the two halves of the brain or the beginning of the development of a conditional discrimination based on the task-eye correlation in the *CR* group. Accordingly, the data of each *S* from the last half of the experiment (Days 17-32) were examined and the referred stimulus determined. This deter-

TABLE I

THE NUMBER OF DAYS IN EXPERIMENT I ON WHICH THERE WAS SIGNIFICANT DIFFERENTIAL LATENCY OF RESPONSE TO THE POSITIVE AND NEGATIVE STIMULI

(The data are presented for eight-day blocks of interocular discrimination.)

Condition	*S*	Eight-day blocks 1	2	3	4
NR	16	1	5	8	8
	7	3	6	8	8
	54	0	7	8	8
	57	0	3	8	8
	Mean	1.0	5.25	8.0	8.0
CR	5	0	1	3	2
	1	0	0	1	5
	6	0	2	3	4
	53	0	1	1	0
	52	1	2	0	4
	Mean	0.2	1.2	1.6	3.0
UR	55	1	1	1	1
	10	1	0	1	2
	15	1	2	2	3
	14	0	0	0	3
	58	0	1	1	5
	Mean	0.6	0.8	1.0	2.8

mination was based simply on a comparison of the average differential-latency scores when vertical was the positive stimulus and when horizonal was the positive stimulus, without regard to significance of differences. The results of an analysis of variance indicated that the comparisons of interest—the main *Groups* effect ($F = 0.36$, $df = 1/8$) and the interaction of *Stimulus-Preference* $\times$ *Groups* ($F = 1.37$, $df = 1/8$)—were not significant. The deliberate classification of stimuli as preferred and non-preferred was reflected in a significant *Stimulus-Preference* effect ($F = 31.59$, $df = 1/8$, $p < 0.01$), while the *Blocks* effect ($F = 1.91$, $df = 1/8$) was not significant. It may be concluded that there is no evidence here of any thing less than complete interocular transfer.

The results for two binocular control *Ss* do not supply evidence for a differential effect of alternating and nonalternating schedules of task-reversals. Neither *S* gave evidence of discrimination between the negative and positive stimuli. Each tended to hit the target indiscriminately in the presence of either stimulus.

Following the last day of interocular discrimination, one *NR S* (7) was placed on a daily reversal-schedule for 20 days; the originally positive stimulus remained positive when the right eye was uncovered, while reversal training was given on days when the left eye was uncovered. Another *S* from the same group (16) remained on the original training schedule for nine additional days and then was changed to a daily reversal schedule for 11 days during which the left eye was always uncovered. The first *S* showed immediate negative transfer from previous learning on reversal-sessions. During the first four reversal-sessions, a greater proportion of responses were made on trials with the negative (formerly positive) stimulus. In the next four reversal-sessions, there was no difference in probability of choice of the two stimuli (or in latency of response). On the final reversal-session, over-all probability of response dropped sharply. Discrimination was maintained on non-reversal sessions until the last session during which over-all response-level fell. The animal failed entirely to hit the target on two subsequent placements in the experimental apparatus. The second *S* (16) showed no evidence of reversal-learning during the six sessions on which the originally negative stimulus was positive. Discrimination on the original task was maintained for a few days following which the over-all probability of response decreased sharply.

Experiment II

Experiment II was essentially a replication of Experiment I with the exception that a simultaneous discrimination was employed to examine the generality of the results with a new method of stimulus-presentation and to a new measure of response. Only those features of the present experiment which differ from those of Experiment I will be described.

Subjects. The *Ss* were 18 goldfish taken from the same stock as in Experiment I.

Apparatus. The enclosure differed from that used in Experiment I only in that there were two targets, about 2½ in. apart, with an In-Line projector located behind each. The apparatus is sketched in Fig. 2. All the events of the experiment were programmed automatically and the responses of the animals were recorded on tape.

Adaptation and preliminary training. After preliminary adaptation and feeder-training, *Ss* were taught to hit the plastic target for food-reward in the single-target apparatus used in Experiment I. Binocular target-training and all subsequent

training were carried out in the double-target apparatus. On any trial, only one target was illuminated (with yellow light), the right on half the trials and the left on the remaining trials. Twenty trials were given each day, for four days, by which time the target-response was well established.

Green-red discrimination constituted the next stage of preliminary training. All *Ss* were trained with green as the positive stimulus because some observations made in Experiment I indicated that *Ss* pretrained to strike a yellow target subsequently tend to prefer red to green. Each trial began with the illumination of both targets. A response to the positive target turned off both target-lights and initiated a reinforcement-cycle; response to the negative target was followed by guidance—that is, both stimuli were turned off and, after a 6-sec. interval (time-out), the positive stimulus alone reappeared. During this stage of training, the animals were given 20 trials per day until a criterion of no more than two errors in a given session was reached, which took an average of five days.

Monocular target-training followed, with 10-20 trials given each day until the target-response was reëstablished—that is, until *S* was hitting the lighted target consistently, and until total response-time for 20 trials was 10 min. or less. From 70-180 trials were required for *Ss* to meet this criterion. During all monocular trials, the feeder-light remained on for 15 sec.

On the basis of their performance during the final two stages of preliminary training, the *Ss* were assigned to the various experimental and control conditions (2 *NR*, 3 *CR*, 4 *UR*, and 1 Nonreversal control). Of the original 18 *Ss*, five had become ill, one was dropped from the experiment following a failure to show any evidence of learning the green-red discrimination, and two failed to meet the time-criterion during monocular training despite prolonged training.

Interocular discrimination. The experiment consisted of 32 daily sessions of 30 trials each. Except for differences already noted, the design was the same as that of Experiment I. The procedure in dealing with errors was changed during the course of the experiment. In the initial sessions, an unlimited correction procedure was employed; that is, when an initial error was made, the stimuli behind both targets were turned off and both reappeared following a 6-sec. time-out. An unlimited number of repetitive errors could be made, and a correct response terminated the trial. The unlimited correction procedure was adopted on the chance that a repetitive-error measure might differentiate the groups. No substantial differences appeared in the first 23 sessions, however, and thereafter the guidance procedure was used.

Results. The results of this experiment are presented in Fig. 4, in which percentage correct choice per day is plotted for each *S*. The appropriate points have been joined to yield separate learning curves for right and left eyes. The curves for the two *NR Ss* rise gradually, and at about the same rate for the right eye as for the left. By contrast, the curves for the *CR* and *UR Ss* remain at or near the chance-level. The curve for the binocular control indicates a level of performance quite comparable to that yielded by the *NR* condition.

The final mean percentage of correct choice, based on the last two

sessions wtih the left eye and the last two sessions with the right eye, is 83.2 for the *NR* group, 50.4 for the *CR* group, and 53.4 for the *UR* group. The results of an analysis of variance of these data showed a significant

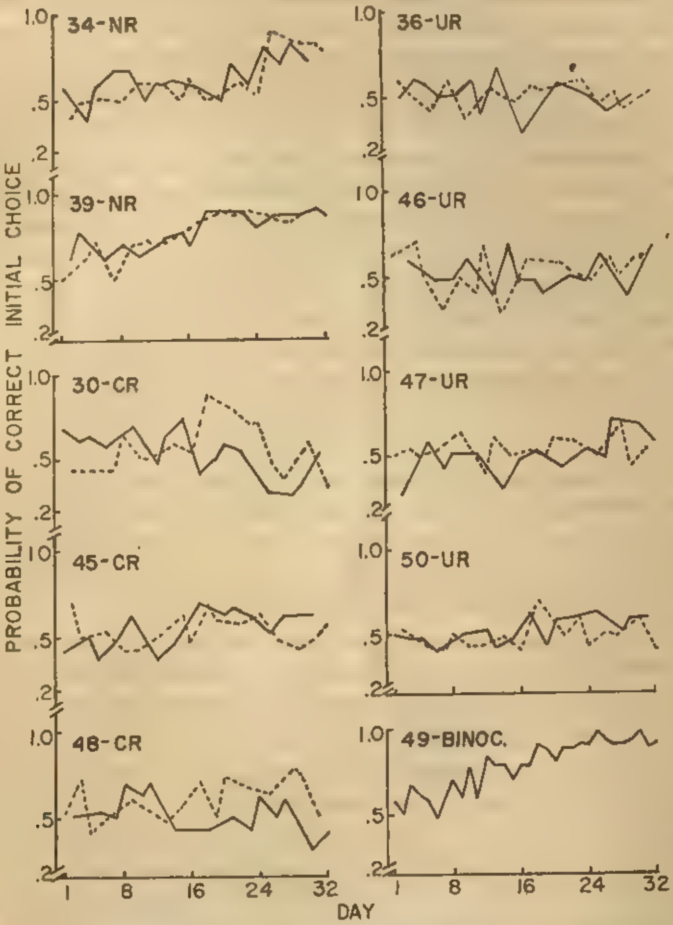

FIG. 4. INDIVIDUAL CURVES SHOWING THE PERCENTAGE CORRECT CHOICE ON EACH DAY OF EXPERIMENT II

(For experimental *Ss* the left-eye data appear as a solid line, the right-eye data as a broken line. Only a single curve is shown for the binocular control.)

Groups effect ($F = 26.9$, $df = 2/6$, $p < 0.01$). A subsequent Scheffé test indicated that the *NR* group differed significantly from the *CR* and *UR* group ($p < 0.01$), but that there was no difference between the *CR*

and *UR* groups. Thus, the results of Experiment II closely parallel those of Experiment I.

After one *NR S* (34) had been trained with horizontal positive for 32 days, the vertical was made positive in left-eye sessions while horizontal remained positive in right-eye sessions. The results indicate negative transfer in right-eye sessions; by the sixth session, performance with each eye had fallen to chance-level. It is possible, of course, that with additional training a conditional discrimination would have developed, but no evidence of functional separation of the two eyes appears here.

The binocular control animal (49) was tested monocularly (left eye) following training to an asymptote of about 90%. Performance in the first monocular session fell to the 57% level, with recovery, in the following three sessions, to 80%, 77%, and 80%, respectively. The blinder then was placed over the right eye, and again performance dropped (to the 60% level) in the first session, but improved in the second and third session (80% and 83%, respectively).

Discussion

The results of these experiments indicate that interocular transfer of a pattern discrimination—at least one of the level of difficulty of the stripe-discrimination used here—is complete in the goldfish when appropriate controls are included for (1) orientational difficulties and specific response-biases associated with monocular vision and the restriction of the visual field, and (2) general learning about the experimental situation which is not specific to the particular discrimination presented. The results suggest that previous failures to find any interocular transfer at all, and reports of less than perfect transfer, reflect only the lack of a suitable procedure.

The highest level of transfer of a visual discrimination in an instrumental situation previously reported was in Schulte's experiment, in which *Ss* were highly trained on one eye prior to testing for interocular transfer.[18] Unfortunately, the *Ss* in that study were preëxposed binocularly to the stimuli to determine stimulus-preferences before monocular training was begun. In the present study, transfer was examined at all stages of learning, and the ambiguity of transfer-scores—which may reflect transfer-effects both specific and general, positive and negative—was eliminated.

[18] Schulte presented data for one discrimination which is similar, though not identical, to the present task. Of three *Ss* trained to discriminate between a tall, narrow, inverted triangle and a horizontal bar, one made a high transfer-score, one showed a drop of more than 20% which may have reflected general debilitation, and one performed at the chance-level (position habit) when tested monocularly with either eye.

The design employed in the present study is, of course, applicable to situations in which more complex responses are used, as in the early maze-studies of McCleary which gave only negative results. If McCleary's suggestion that failure of interocular transfer was due to a failure of new visuomotor learning to transfer is correct, then it follows that, if *Ss* are pretrained to include the relevant response in their (monocular) repretoires, there should be complete interocular transfer of a subsequently acquired discrimination.

In the present study, there was no sign of conditional discrimination when task and eye-usage were correlated from the beginning of training (*CR* condition). Nor did a conditional discrimination develop in the case of two *Ss* (7 in Experiment I; 34 in Experiment II) trained to criterion under *NR* conditions and subsequently trained on the reversed discrimination when one eye was used but not when the other eye was used. In one case (7), discrimination was maintained for a time in the original eye, while there was no sign of reversal-learning when the other eye was used (*e.g.* the original preference persisted). In the second case (34), performance fell to the chance-level when either eye was used. The latter result is in agreement with Schulte's observation, as is the final refusal to the first *S* to perform. It is, of course, possible that with an easier discrimination, and with additional training, a conditional discrimination might be developed.

The results for another *S* (49) suggest that response-biases are associated with monocular performance, and that new response-learning takes place when a blinder is transferred from one eye to the other. This animal had been performing at a high level during binocular visual discrimination. A sharp drop in performance, with subsequent recovery, (which was not complete after four monocular testing sessions) occurred when a blinder was placed on one eye and when it was transferred to the other eye.

The question arises as to which neural paths might mediate interocular pattern-transfer in the goldfish. Bernstein demonstrated that interocular transfer of a brightness-discrimination and a hue-discrimination was not permanently interfered with by unilateral (ipsilateral or contralateral) or by bilateral forebrain ablations.[19] Ablations were made following training, and a cardiac-deceleration index of discrimination was used. Although Bernstein concluded that the forebrain commissures are not necessary for interocular transfer, the possibility remains that ablations made before original training would prevent subsequent interocular transfer. It also is

[19] J. J. Bernstein, Role of the telencephalon in color vision in the fish, *Exper. Neurol.*, 6, 1962, 173-185.

possible that pattern-discriminations may yield a different outcome. Bernstein noted that Schade and Weiler failed to record evoked potentials from the optic tectum of the goldfish upon photic stimulation of the eye, and therefore eliminated the midbrain as a source of crucial commissures for interocular transfer. He suggested that one or more of the diencephalic commissures (there are three such commissures crossing in the diencephalon) subserves interocular transfer.

Summary

Two experiments designed to test for and assess the completeness of interocular transfer in the goldfish are reported. In each, the *Ss* were given monocular target-practice prior to interocular discriminative training in order to eliminate transfer of motor habits as a source of variance in the assessment of the amount of visual transfer.

All the *Ss* were taught to discriminate between a horizontal and a vertical stripe. In Experiment I, a single-target apparatus was used and a latency-measure of response was employed. In Experiment II, a two-target apparatus was used and a choice-measure was employed. In both experiments, Non-reversal (*NR*) *Ss* were trained with the same stimulus positive regardless of which eye was used during a daily session. Correlated-reversal (*CR*) *Ss* were trained with one stimulus positive when the right eye was used and the other stimulus positive when the left eye was used. For Uncorrelated-Reversal (*UR*) *Ss*, the positive stimulus in any reversal was uncorrelated with the eye in use. The *Ss* which learned the discrimination binocularly, also were studied.

The superiority of the *NR* performance over the *CR* performance is interpreted to mean that some transfer is present, since, if functional separation of the eyes were complete, right- and left-eye learning would be independent, and it would not matter whether the same or opposite discriminations were being learned by the two eyes. The lack of significant differences between *CR* and *UR* performances is interpreted to mean that transfer was complete, since any degree of functional separation of the eyes would be expected to result in superior learning when a given discrimination is consistently associated with the use of one eye.

There was no evidence for the establishment of a conditional discrimination in *Ss* of Group *CR* in either experiment. In fact, neither the *CR* nor the *UR* animals gave any evidence of learning in either experiment, while the *NR* animals showed significant discrimination.

CLUSTERING AS A FUNCTION OF ASSOCIATIVE COMMONALITY *

By FREDERICK R. FOSMIRE, University of Oregon

Jenkins, summarizing the work of the Minnesota group on language-stylistics, concludes that associative commonality is a concept of limited usefulness, which refers only to intraverbal habits.[1] Apparently, the later research of Jenkins did not support the initial impression that simple commonality of word-association may index groups of persons who are reliably different across a great variety of perceptual, cognitive, and adjustive, as well as linguistic, aspects.

In one of the early studies, for example, Peterson and Jenkins studied two *Ss* intensively: one high in commonality, *i.e.* giving many popular or conventional associations, who was scored *H;* and the other, low in commonality, *i.e.* giving few popular or conventional associations, who was scored *L*.[2] After numerous experiments with these *Ss*, they found that *H* was more "consistent," "redundant," and "predictable," and that *L* was more "inconsistent," "non-redundant," and "unpredictable." They concluded from these results that "conformance to cultural responses in a test of word-association does not represent performance on one specific, isolated test, but rather is indicative of a broad, generalized trait of commonality of verbal conformance."[3]

This study, then, is in part a follow-up of the work of Peterson and Jenkins to determine whether *Ss* classified on the basis of high (*H*) and low (*L*) commonality in tests of free association manifest associative differences in other than discrete associative tasts, and to explore the heuristic value of cluster-analysis. The cluster-analyses reported here are

* Received for publication June 4, 1963. This study was financed in part by a grant-in-aid from the Graduate School of the University of Oregon and by NIMH grants H-3841(A) and M-4978. My thanks are due Miss Joyce Thomas and Mr. H. Edward Tryk for their assistance.

[1] James J. Jenkins, Commonality of association as an indicator of more general patterns of verbal behavior, in T. A. Sebeok (ed.), *Style in Language*, 1960, 307-329.

[2] M. S. Peterson and J. J. Jenkins, Word-association phenomena at the individual level: A pair of case studies, *Onr Technical Report No. 16*, Univ. Minn., 1957, 46-47.

[3] *Idem*, 46.

similar to, but differ in a number of respects from, those employed by Jenkins and his colleagues.[4]

Method and procedure. Serial association was chosen as the method of study, since it possesses many of the characteristics of a projective technique without forcing *S* to project. One *S* might give ordinary associations through all cycles, the words occurring most frequently in his total production of responses being those most common in the English language. Another *S* might quickly 'collapse' onto a few idiosyncratic associations which have personal significance.

The principal dependent variables in this research are: (1) associative clusters; and (2) focus-words. An 'associative cluster' is a set of words related by manifest association, *e.g.* 'love' and 'hate' comprise a cluster for a given *S* if either, given as a stimulus-word, evokes only the other word. Figs. 1 and 2 are examples of clusters.

A focus-word is particularly rich in association. It may elicit a number of different responses, it may be elicited by a number of different stimulus-words, or both. Extending Noble's concept of meaningfulness (*m*) to include backward associations, focus-words might be thought of as being more meaningful than other words in a cluster.[5] A word having five or more associates was arbitrarily defined as a focus-word.

On the assumption that *S*s classified as *H* are more constrained by a stimulus-word to make a particular high-strength response, and that the *Ls* have a greater number of somewhat weaker associations to a given stimulus-word, the *H*s and *L*s were expected to produce associative clusters that differed between groups in size and complexity, *i.e.* the *H*s were expected to have many two- and three-word clusters and the *Ls* should tend to move toward a few personally significant, conceptually related words and produce large complex clusters. If size and complexity of associative clusters are regarded as structural indices, the content of the clusters should be different also, but in ways that were not specified *a priori* other than that the words of the *H*s might be more neutral and impersonal than those produced by the *Ls*.[6]

(1) Subjects. The Kent-Rosanoff word-association test was administered to more than 400 unselected undergraduate students. From this number, 4 men and 5 women, who gave 14 or fewer popular words, were selected to form the low commonality group (*Ls*); and 2 men and 5 women, who gave 47 or more popular words, to form the high commonality group (*H*s).

(2) Task. A new list of 44 stimulus-words was constructed for use in the serial associative task (see Table I). Most of the words in the Kent-Rosanoff list are common words, as indexed by their Thorndike-Lorge values,[7] and it may be that common stimulus-words tend to elicit common words as responses. The new list

---

[4] Peterson and Jenkins, *op. cit.*, 1-49; J. J. Jenkins and W. A. Russell, Associative clustering and commonality of responses in the Kent-Rosanoff test, *Our Technical Report No. 3,* University Minnesota.

[5] C. E. Noble, An analysis of meaning, *Psychol. Rev.*, 59, 1952, 421-430.

[6] Peterson and Jenkins (*op. cit.*, 39) contrast the expository style of the two groups of *S*s, describing *H*'s writing as "flat, dull, and trite."

[7] Irving Lorge and E. L. Thorndike, *A Semantic Count of English Words,* 1938.

is called the *Variable Frequency List (VFL)*, as it samples equally and randomly all levels of commonness in the Thorndike-Lorge tables. The only constraint placed upon selection was that the words be nouns[8] and must not be considered archaic or vulgar.[9] Although the *Ss* were not asked if they knew all of the words, the nature of their responses suggests that they had some idea of their approximate meanings.

A modified procedure of serial association was followed. First, the prepared list of stimulus-words (S_1) was given the *Ss*; then the words given in response (the R_1-words) were used as stimulus-words (S_2); then the words given in response (the R_2-words) were used as stimulus-words (S_3); and so on through five cycles, as indicated in the following schema: $S_1 \rightarrow R_1$; R_1-$S_2 \rightarrow R_2$; R_2-$S_3 \rightarrow R_3$; R_3-$S_4 \rightarrow R_4$; R_4-$S_5 \rightarrow R_5$. A 3-min. break was taken between successive cycles. The *Ss* were instructed to give the first word that "came into mind" after hearing the stimulus-word.

TABLE I

THE VARIABLE FREQUENCY-LIST OF STIMULUS-WORDS

1. Daffodil	12. Apocalypse	23. Typist	34. Genesis
2. Mongol	13. Anesthetic	24. Survival	35. Tornado
3. Execution	14. Umpire	25. Overdue	36. Raid
4. Scenery	15. Brogue	26. Gristle	37. Imprint
5. Sawdust	16. Sunshade	27. Coolie	38. Elixir
6. Shawl	17. Chore	28. Seduction	39. Nightmare
7. Sepulchre	18. Seeker	29. Pitfall	40. Disability
8. Jade	19. Elector	30. Brigand	41. Invader
9. Exponent	20. Mandate	31. Hearse	42. Braggart
10. Ulcer	21. Chicanery	32. Aspen	43. Forerunner
11. Middleman	22. Civilian	33. Analogy	44. Sun

Results. Associative clusters were drafted for every *S*. Figs. 1 and 2 are examples of clusters; Fig. 1 for an *L* and Fig. 2 for an *H*.

(1) First analysis. In the first analysis of the data, the *H*s and *L*s were compared on the following points: (a) number of separate clusters; (b) number of items in the largest cluster; (c) number of items in the smallest cluster; and (d) total number of different words produced.

Table II gives the means and variances for each group on measures (a), (b), (c), and (d), and the *t*s between the means for each of these measures. The only non-significant difference between the groups is for (d), suggesting that the *Ss* do not differ greatly in size of vocabulary. Students' *t* for (c) is significant at the 5% level, for (a) at the 1% level, and for (b) at the 0.1% level. Very clearly, then, the *H*s produce more clusters containing fewer words per cluster. Obviously, these measures are not independent. The comparisons between number of items in the

[8] One adjective was, however, included.

[9] As defined in Webster's *New Collegiate Dictionary.*

largest and smallest cluster are presented for illustrative purposes, not to exaggerate the difference between the two groups in kind or style of clustering.

One might, in interpreting the raw data in (d), think of the range of values theoretically attainable. If *S* gave the same response to all

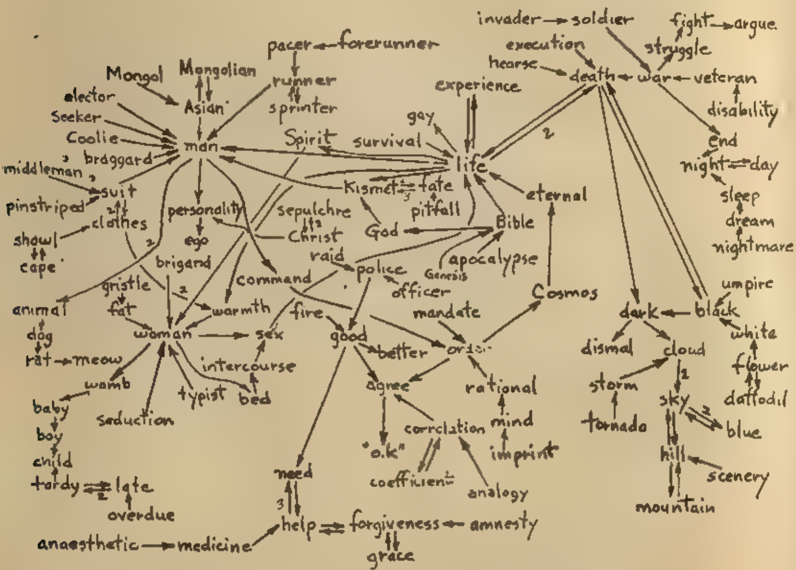

FIG. 1 AN EXAMPLE OF A COMPLEX CLUSTER PRODUCED BY AN *L*

2
pitfall→danger⇄harm
2

FIG. 2 AN EXAMPLE OF A SIMPLE CLUSTER PRODUCED BY AN *H*

stimuli, the total size of his cluster would be 44 + 1, or 45. At the other extreme, if *S*s never gave the same response twice, responding with a different word to every stimulus for five cycles, his total would be 44 + (5 × 44), or 264 words.[10]

[10] To the extent that the Ohio State University Psychological Examination measures vocabulary-size, we have an estimate of the relationship between word-association commonality and vocabulary. The mean stanines for another sample of 14 men and 12 women are 5.5 for the *H*s and 6.2 for the *L*s. Although the *L*s seem to have larger vocabularies than the *H*s, the stanine-difference between the sexes is greater than the difference between *H*s and *L*s.

(2) *Second analysis.* It was anticipated that clusters produced by the *Ls* would be more revealing of feelings and attitudes than those produced by the *Hs*. A method of testing this idea had not, however, been devised before inspection of the clusters. *E*'s early impression was that words having a large number of associates in a cluster were different in affective tone between the two groups. Would a group of judges agree that focus-words produced by the *Ls* were more neutral and impersonal than those produced by the *Hs*?

Method and procedure. Sixteen judges, 6 men and 10 women, were asked to rate the affective tone of the focus-words on a seven-point scale. They did not know how the words were selected.

Results. The *H*- and *L*-groups each produced 23 different focus-words,

TABLE II

COMPARISON OF ASSOCIATIVE CLUSTERS PRODUCED BY THE *Hs* AND THE *Ls*

Ss		(a) No. of clusters	Number of items in (b) largest cluster	Number of items in (c) smallest cluster	(a) Number of different words
Hs	mean	31.29	23.29	2.71	161.29
	σ	5.09	13.13	1.11	45.90
Ls	mean	19.44	66.78	3.67	185.44
	σ	11.01	30.39	1.25	50.45

of which some were given by more than one *S*, resulting in 32 focus-words for each group. The focus-words are given in Table III, which shows the mean rating of affective tone for each '1' as being 'extremely unpleasant' and '7' 'extremely pleasant.' The distribution of the affective ratings is shown in Fig. 3. The *Ls* apparently tend to produce focus-words of definite affective tone, either pleasant or unpleasant, avoiding bland words. The *Hs* have a marked tendency to produce pleasant focus-words. In spite of the bimodal distribution of ratings of focus-words produced by the *Ls*, and the greater range of affectivity at both ends of the distribution, a Mann-Whitney U to Z transformation shows the affectivity ratings to be reliably different ($p < 0.03$).

The analysis reported in Fig. 3 and Table III is based on the affectivity-ratings of individual focus-words. Since both *H*- and *L*-groups produced 23 different focus-words, and the *Ls* produced fewer clusters (see Table II), it follows that the *Ls* produced a larger number of focus-words per cluster. If clusters tend to have a single over-all affective tone, one or two large clusters, each with several focus-words, could be responsible for most of the difference in affectivity ratings.

If, however, the affectivity ratings of focus-words within a single cluster are averaged, thus yielding a median affectivity-rating for each *cluster*, a comparison between the *H*s and the *L*s on cluster-affectivity would not be

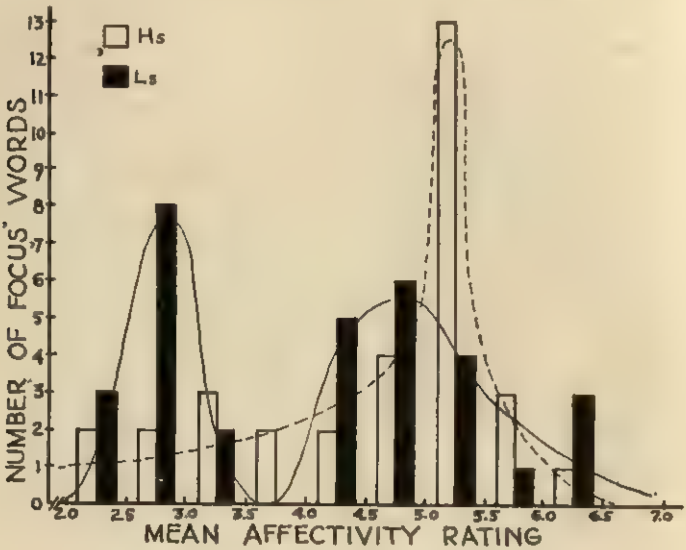

FIG. 3 JUDGES' RATINGS OF FOCUS WORDS PRODUCED BY *Hs* AND *Ls*

TABLE III

AFFECTIVE TONE OF FOCUS-WORDS

Focus-word	Mean rating	Sigma	Focus-word	Mean rating	Sigma	Focus-word	Mean rating	Sigma
1. kill	2.3	2.27	16. yellow	4.3	1.13	30. food	5.3	1.50
2. confusion	2.4	1.46	17. red	4.3	1.62	31. tree	5.4	1.32
3. hurt	2.5	2.16	18. profit	4.3	1.58	32. heat	5.4	1.41
4. mathematics	2.8	1.91	19. rain	4.3	2.12	33. grass	5.5	1.27
5. dead	2.9	2.08	20. industry	4.4	1.20	34. green	5.5	1.32
6. fight	2.9	2.09	21. work	4.7	1.67	35. leaves	5.5	1.37
7. black	2.9	1.19	22. woman	4.7	1.84	36. person	5.5	1.41
8. grave	2.9	2.00	23. law	4.9	1.23	37. man	5.6	1.20
9. death	2.9	1.96	24. fair	4.9	1.41	38. capable	5.6	1.30
10. gook	3.1	1.63	25. study	4.9	1.46	39. soft	5.7	.79
11. hard	3.3	1.62	26. impression	5.1	1.20	40. book	5.9	1.03
12. nothing	3.3	1.69	27. tall	5.1	1.18	41. sun	6.0	.97
13. opponent	3.6	1.59	28. game	5.3	1.01	42. sleep	6.1	.99
14. business	4.0	1.75	29. dog	5.3	1.29	43. life	6.3	.87
15. soldier	4.2	1.59				44. love	6.5	1.64

influenced by any tendency for words within a cluster to be similar in affective tone. Such a comparison yields a Mann-Whitney U that is significant beyond the 1% level. Thus, the difference in affective tone seems not to be artifactual.

(3) Third analysis. Postman has shown that there are many more words of pleasant than unpleasant connotation in the English language and

argues that the affective tone may determine how frequently a word is used in the general population.[11] He recognized, however, that the opposite causal sequence might hold: "that familiarity resulting from frequency may be in itself a source of positive value."[12] Johnson, Thomson, and Frincke confirmed that pleasant words are more common than unpleasant ones (of the order of 2:1, they suggest) and found a significant correlation between frequency of usage (Thorndike-Lorge count) and rated 'goodness' of words.[13] These experimenters also demonstrated that *S*s' evaluation of words increases as a function of familiarity.

With respect to our experiment on associative clustering, then, *E* was left with the possibility that the *H*s produce only very common words as focus-words, hence would tend to give more 'pleasant' words. If the principal difference between the *H*s and *L*s is that the *H*s give common associations and the *L*s give uncommon ones, the difference in affectivity of focus-words could be an artifact of word-familiarity.

Method. E returned to the Thorndike-Lorge tables and found the value of each focus-word that was contained in the tables. Median values for the two groups of *S*s were computed.

Results. There is no difference in median Thorndike-Lorge value between the two groups of focus-words. Thus, the hypothesis that the difference in affective tone of focus-words is an artifact of commonness can be rejected.

Discussion and Conclusions

If we think of the size and complexity of clusters as representing some *structural* aspect and affectivity of the focus-words as representing one *content* aspect of cognitive organization, then we have additional evidence that commonality of word-association is a useful index of cognitive habits. The following interpretation of these experiments emphasizes the importance of three observations about *H*- and *L*-*S*s: the difference between the two groups in consistency and redundancy, and the manner in which the groups of *H*- and *L*-*S*s maintain their separate identities on a variety of tasks mediated by language.

The greater consistency of the *H*s seems to be fairly well established. Peterson and Jenkins stress the *H*s' greater consistency across tasks, referring

---

[11] Leo Postman, The experimental analyses of motivational factors in perception, in J. S. Brown (ed.), *Current Theory and Research in Motivation,* 1953, 59-108.

[12] Postman, *op. cit.,* 68.

[13] R. C. Johnson, C. W. Thomson, and Gerald Frincke, Word values, word frequency, and visual duration threshold, *Psychol. Rev.,* 67, 1960, 332-342.

to them as "consistent," "redundant," and "predictable."[14] With respect to associative clustering in a *recall* test, Jenkins and Russell report, "In general, *H* used a few words in tightly organized groups which tended to be stable both within the experiment and from time to time while *L* used many words in a complex network of interrelationships which tended to shift from one moment to the next."[15]

Further developing some ideas borrowed from Newcomb and from Coombs, Runkel suggests that response-consistency is most likely when an individual imposes simple order on a multidimensional set of responses.[16]

The various findings are consistent with the hypothesis that, although different persons give differential weightings to the attributes of a stimulus (or its referent, in the case of a stimulus-word), the *H*s are more likely than the *L*s to attend characteristically to only a few attributes, and to assign differential importance to the attributes of a particular stimulus in a stable, hierarchical order. Thus, on a test-retest measure of the reliability of word-association, the *H*s should be more consistent than the *L*s. When *E* tested this hypothesis, the data supported it, both in terms of numbers of populars and of numbers of exact repetitions. Similarly, to push Runkel's hypothesis farther, small, independent, associative clusters might be thought of as representing relatively unidimensional response-spaces, and large, complex, clusters as representing an aggregate of partially overlapping multidimensional spaces.

If one can demonstrate that the associative response to a stimulus-word results from the same constraints that operate in connected discourse, then differences in redundancy can be explained in the same terms employed to account for differential consistency in discrete association. Jenkins and Russell, although not writing in the context of relating cluster-characteristics or commonality to redundancy, state: "It has been demonstrated that when an *S* produces a word, the effect of this word upon his subsequent production is analogous to the effect that such a word would have if spoken by another person as a stimulus-word in a free association-test."[17]

Summary

Interested in word-association commonality as a way of identifying groups of persons who demonstrate stable linguistic differences, *E* investi-

[14] Peterson and Jenkins, *op. cit.*, 46-47.
[15] Peterson and Jenkins, *op. cit.*, 9.
[16] P. J. Runkel, Cognitive similarity in facilitating communication, *Sociometry*, 19, 1956, 178-191.
[17] Jenkins and Russell, *op. cit.*, 6.

gated the serial associative task and a method of analysis of associative clusters to determine their heuristic value to research in thought and language. In Analysis I, the structural properties of the associative clusters obtained from two groups of *S*s, who represented opposite ends of the range of commonality in word-association, were compared. The *S*s high in commonality (*H*) produce many small (two-, three-, and four-word) clusters, while the *S*s low in commonality (*L*) tend to give large, complexly interrelated clusters.

Because words associated with several other words should be more meaningful to *S* than words associated with only one or two other words, a word having five or more associates was defined as a focus-word. In Analysis II, an independent group of judges rated the focus-words for pleasantness-unpleasantness. The two groups of *S*s differed significantly: the *H*s produced more neutral and pleasant words, while the *L*s gave focus-words that were more polar and, on the average, less pleasant in affective tone.

Analysis III tested the hypothesis that the focus-words of the *L*s were more likely to be perceived as unpleasant because of their uncommonness. Comparison of the lists of focus-words obtained from each group revealed a non-significant difference in median Thorndike-Lorge value.

Size and complexity of the clusters may be interpreted in terms of cognitive complexity, more specifically to differential tendencies to respond as if simple order had been imposed on a multidimensional response-space. Not only does this formulation account for the differences between the *H*s and the *L*s in associative clustering, it is consistent with observations about their differences in consistency and redundancy.

A COMPARISON OF KEY-PECKING WITH AN INGESTIVE TECHNIQUE FOR THE STUDY OF DISCRIMINATIVE LEARNING IN PIGEONS

By ELIZABETH M. STEARNS and M. E. BITTERMAN, Bryn Mawr College

The results of some early experiments with a variety of species suggest that discriminative learning is particularly rapid when the discriminanda are visually differentiated food-objects which *S* is invited to ingest, the negative objects being saturated with quinine, or glued to the substrate, or otherwise rendered unconsumable.[1] As yet, however, there has been no satisfactory comparison of ingestive techniques with the more conventional techniques, which involve discriminanda-manipulanda other than food-objects and indicator-responses other than ingestion. It is clear from a number of conventional experiments, mostly with monkeys, that close spatial contiguity of discriminandum with manipulandum, and of both with reward (or nonreward), are favorable conditions for learning,[2] but whether ingestive techniques, which maximize contiguity, offer any unique advantages has not been determined. We report here an experiment with pigons in which the effectiveness of a new ingestive technique was compared with that of key-pecking.

METHOD

Magazine-discrimination. The ingestive technique offered the animal a series of choices between two adjacent Gerbrands grain-magazines, one lighted blue and the

* Received for publication June 20, 1963. This research was supported by the Office of Naval Research under terms of contract Nonr 2829(01).

[1] J. Reighard, An experimental field study of warning coloration in coral reef fishes, *Publ. Carnegie Inst., Wash.*, 103, 1908, 257-325; David Katz and Géza Révész, Experimentell-psychologische Untersuchungen mit Hühnern, *Z. Psychol.*, 50, 1909, 93-116; W. T. Shepherd, Some mental processes of the rhesus monkey, *Psychol. Monogr.*, 12, 1910, (No. 52), 1-61.

[2] M. E. Jarvik, Discrimination of colored food and food signs by primates, *J. comp. physiol. Psychol.*, 46, 1953, 390-392; Jerome Wodinsky, M. A. Varley, and M. E. Bitterman, Situational determinants of the relative difficulty of simultaneous and successive discrimination, *ibid.*, 47, 1954, 337-340; G. E. McClearn and H. F. Harlow, The effect of spatial contiguity on discrimination learning by rhesus monkeys, *ibid.*, 47, 1956, 391-394; Jarvik, Simple color discrimination in chimpanzees: Effect of varying contiguity between cue and incentive, *ibid.*, 49, 1956, 492-495; J. V. Murphy and R. E. Miller, Effect of the spatial relationship between cue, reward, and response in simple discrimination learning, *J. exp. Psychol.*, 56, 1958, 26-31; D. R. Meyer, V. J. Polidora, and D. G. McConnell, Effects of spatial S-R contiguity and response delay upon discriminative performances by monkeys, *J. comp. physiol. Psychol.*, 54, 1961, 175-177.

other lighted green. (Christmas-tree lamps were mounted in the magazines in such a way that their interiors could be flooded with colored light.) Choice was detected by photocells, one recessed in the lower lip of each magazine, which were shadowed by the insertion of *S*'s head; illumination of the cells was provided by the colored Christmas-tree lamps in the magazines. The two magazines were mounted side by side in a panel which also contained a single pigeon-key displayed behind a circular opening (1 in. in diameter and centered 3.5 in. above the magazine-apertures). The panel, which was set into a picnic-chest, served as one wall of the animal's chamber.

Each trial began—after a 6-sec. intertrial interval (*ITI*) in darkness—with the illumination of the center-key (*C-K*) by a white lamp. A single peck at the *C-K* darkened it and turned on the colored lights of the magazines. If the correct magazine was chosen, the lights stayed on for 3 sec. when grain was available. An incorrect choice turned off the magazine-lights and initiated a 6-sec. time-out (*T-O*) in darkness, after which the *C-K* was illuminated once more. Some animals were trained with *guidance* and others with *limited correction*. In guidance, response to the *C-K* after *T-O* illuminated the correct magazine alone, and the animal was permitted to eat from it. In limited correction, both magazines were lighted by response to the *C-K* after *T-O*, and the animals were required to make another choice; after the tenth repetitive error on any trial (that is, after the eleventh *T-O*), there was a guided reinforcement and the trial terminated.

On 20 of each day's 40 trials, the right-hand magazine was blue and the left-hand magazine was green; on the remaining trials, the magazine on the right was green and the magazine on the left was blue. The two kinds of trial were scheduled in selected Gellermann-orders. Of the 12 *Ss* studied in this portion of the experiment (experimentally naïve White Carneaux cocks maintained at 75% of their free-feeding weights), 6 were trained in a series of *daily reversals* (3 with guidance and 3 with correction), and the remaining 6 in a series of *criterion-reversals* (3 with guidance and 3 with correction). In the criterion-training, the correct color was reversed for a given animal whenever, in a given session, it made 34 or more correct choices, with at least 17 correct choices in the last 20 trials of the session. For the daily-reversal animals, the correct color was reversed after each session, irrespective of performance. Half the daily animals and half the criterion-animals began with blue positive, the rest with green positive. The entire experiment was programmed automatically, the main events being recorded on a Gerbrands six-pen polygraph.

Standard key-discrimination. The data from the discrimination of magazines were compared with those from the discrimination of keys previously collected in this laboratory—for daily reversal by Bullock and Bitterman,[3] and for criterion-reversal by Stiz.[4] The key-apparatus was like that used for the discrimination of magazines, except that the panel offered a choice between two pigeon-keys illuminated with colored lamps, and correct choice was rewarded by access to a single grain-magazine illuminated with white light. The key-data were not collected under conditions exactly comparable in all other respects to those under which

[3] D. H. Bullock and M. E. Bitterman, Habit reversal in the pigeon, *J. comp. physiol. Psychol.*, 55, 1962, 958-962.

[4] Virginia Sitz, unpublished Honors paper, Bryn Mawr College, 1962.

the magazine-data were collected. Sitz used White King pigeons, and she trained them all with pure correction (that is, with no systematic limitation on the number of repetitive errors). Bullock and Bitterman trained their animals at 80% instead of 75% weight, and with red vs. yellow instead of blue vs. green lights. In neither of the previous experiments was a *C-K* used, the intertrial interval (*ITI*) being followed immediately by the illumination of the two keys. (The *C-K* was used in the magazine-discrimination to rule out the possibility that the animal would have its head in one of the magazines when the magazine-lights came on.) The fact that the conditions for the key-reversals and for the magazine-reversals were not identical meant that it would be impossible to make an exact comparison of the efficacy of the two techniques as such, but the function of the key-data was only to provide a baseline of reasonably good contemporary practice in terms of which to decide whether the magazine-technique had any special promise. When the rate of reversal-learning proved to be much greater for the magazine-discrimination than for the standard key-discrimination, an effort was made to find out why.

Control key-discrimination. A new group of six pigeons (experimentally naïve White Carneaux cocks reduced to 75% of their free-feeding weights) was trained in a key-discrimination *with the magazine-light the same color as the positive stimulus.* This control procedure was designed to test the hypothesis that what was responsible for the superiority of the magazine-animals was not any of the several procedural variations already noted, nor the difference in the nature of the indicator-response, but the exposure to the positive stimulus during reinforcement which the magazine-technique afforded and the key-technique did not. The conditions for the control animals were the same in other respects as for the magazine-animals—white *C-K*, blue vs. green discrimination, with half the animals on guidance and half on limited correction—except that the grain-tray was presented for 4 sec. (instead of 3 sec.) in the control condition. (The reinforcement-time for the standard key-animals also was 4 sec.) A 3-sec. presentation was set for the magazine-animals in the hope of equating actual eating time, since the head of a magazine-animal was already in the aperture of the correct magazine when the tray was raised. Actually, however, 3 sec. proved somewhat too long, since some of the animals tended to gain too much weight on the 40 reinforcements earned each day, and the time was gradually reduced still further (by different amounts for different birds) to an average of 2.4 sec.

The control animals were trained first in a series of 12 criterion-reversals, after which they were shifted to daily reversal, and then some further changes in the conditions of training were introduced. In one case, the magazine-light was white; in another, the magazine-light was the color of the *in*correct key; in still another, the colors to be discriminated were changed from blue and green to red and yellow (really a yellowish orange) or to blue and yellow.

Results

In Fig. 1, the daily-reversal performance of the magazine-animals (correction and guidance) is compared with the performance of two standard key-animals trained (guidance) by Bullock and Bitterman.[5] The curves

[5] Bullock and Bitterman, *op. cit.*, 960 (Fig. 2, Pigeons *G* and *H*).

show a striking superiority of the magazine-technique. All 6 of the magazine-animals improved rapidly in the 20 sessions, while neither of the key-animals showed any substantial improvement in those sessions. The magazine-animals trained with guidance had considerably more difficulty in the first reversal (Day 2) than did those trained with correction, but the performance of the two groups was essentially the same thereafter. The divergence of the two magazine-curves in the last five or six sessions reflects the inexplicably erratic performance of one of the animals in the

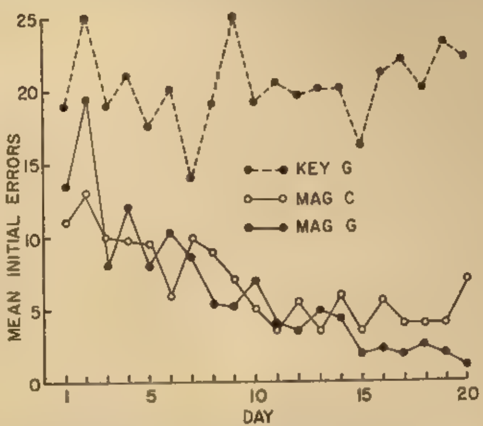

FIG. 1. PERFORMANCE OF STANDARD KEY-ANIMALS AND MAGAZINE-ANIMALS IN A SERIES OF DAILY REVERSALS.
The key-animals were trained with guidance (*G*). Some of the magazine-animals were trained with guidance and others with correction (*C*).

correction group. In all other respects, variability within groups was small, and the three mean curves well represent the behavior of the individual animals in the three groups.

In Fig. 2, the criterion-reversal performance of the magazine-animals (correction only) is compared with the performance of Sitz's standard key-animals and with the performance of the control key-animals (correction only). The curves are plotted in terms of initial errors alone, but the curves of the repetitive-errors have the same shape as do those for initial errors, rising at first, and then declining progressively, with the magazine- and control-groups making substantially fewer errors throughout. The mean curves are, in general, quite representative of the individual performances, although there was one deviant animal in the magazine-group,

whose relatively poor performance in the early reversals was primarily responsible for the higher mean error-level of the magazine- as compared with the control-group. Even so, however, there was no overlap at all between the magazine- and standard key-groups, or between the control- and standard key-groups. Between the magazine- and control-groups, of course, there was considerable overlap; an examination of the two sets of individual functions reveals no systematic difference between them.

The magazine- and control-groups trained with guidance made more errors in the early reversals than did the corresponding correction-groups,

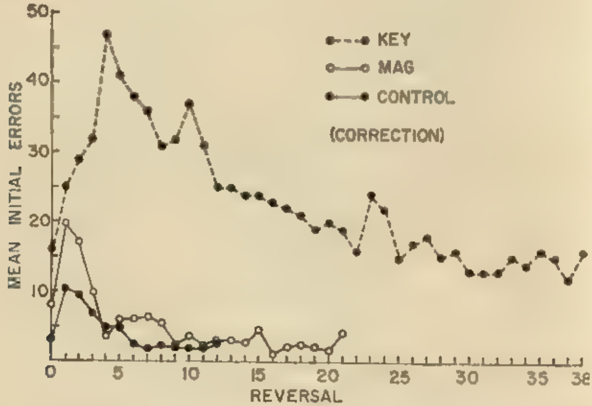

FIG. 2. PERFORMANCE OF STANDARD KEY-ANIMALS, CONTROL KEY-ANIMALS, AND MAGAZINE-ANIMALS IN A SERIES OF CRITERION-REVERSALS
All animals were trained with correction

although the difference soon disappeared. For the pooled magazine- and control-groups (combined to permit statistical treatment), correction produced significantly fewer errors than did guidance in the first reversal ($p < 0.02$) by Wilcoxon's non-parametric test for unpaired replicates. Criterion-reversal functions for the magazine-animals trained with correction and guidance are plotted in Fig. 3. The curves show the early difference between correction and guidance, as well as its rapid disappearance.

Do the magazine- and control-techniques yield lower asymptotic error-levels than the standard key-technique, as well as greater rates of improvement? Probably not. Three of the four Bullock-Bitterman animals eventually (in about 70 days) reached the same level as the magazine-animals trained with daily reversals. Three of the five animals trained

by Miss Sitz eventually (after 32-50 reversals) reached the same level as the magazine- and control-animals trained with criterion-reversals, and the other two were continuing to improve when their training was terminated. It seems safe to conclude that the superiority of the magazine- and control-techniques is limited to the rate at which the reversal-capability is acquired.

After they had completed 12 criterion-reversals, all six control animals were trained on a series of daily reversals with guidance. While Subgroup *C*-1 (consisting of two of the former correction-animals and one of the former guidance-animals) continued under the control condition (with the magazine-light the same color as the positive key), Subgroup *C*-2 (consisting of the remaining three animals) was shifted after two days to the

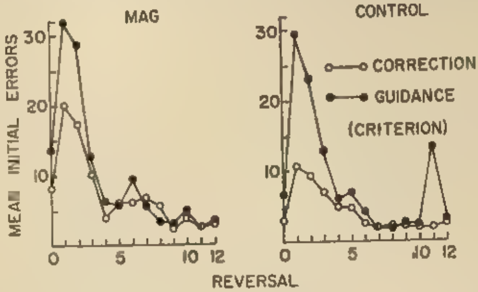

FIG. 3. PERFORMANCE OF MAGAZINE- AND CONTROL-ANIMALS TRAINED WITH CORRECTION AND WITH GUIDANCE IN A SERIES OF CRITERION-REVERSALS

standard key-discrimination (magazine-light always white); with the exception of magazine-light color, the training of the two subgroups remained the same. As Fig. 4 shows, the white magazine-light produced marked and stable disruption of performance in Subgroup *C*-2, but the effect was not manifested until the *second* day of the standard procedure. The phenomenon of delayed disruption was noted again in these animals when, after having been returned in the fifteenth session to the control condition and continued on that condition for the next seven sessions, they were shifted, beginning with the twenty-third session, from the blue vs. green to red vs. yellow discrimination (with magazine-light the same color as the positive key). The disrupting effect of this change, too, appeared for the first time in the *second* session with the changed procedure. After 11 red-yellow sessions during which there was some gradual improvement in performance, the *C*-2 animals were shifted to blue-yellow, which produced an immediate return of performance to the control baseline. In the thirty-eighth session, the white-magazine condition was intro-

duced again, and again there was delayed disruption of performance, although the disruption was less extensive than before, and there was every indication (when it became necessary to terminate the experiment) that the control level soon would be achieved.

In the seventeenth session, the C-1 animals were shifted from the control condition to a reverse-magazine condition (magazine-light the same color as the *negative* key). The resulting disruption of performance was immediate and much more acute than in the white-magazine condition; while the white magazine produced a chance-level of error, the reverse

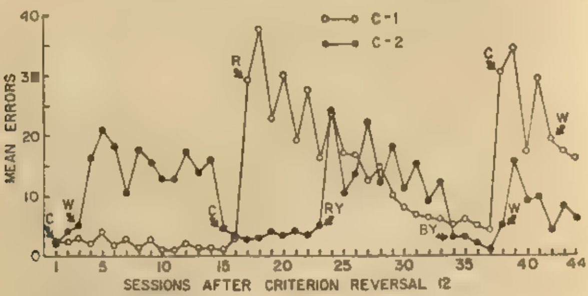

FIG. 4. PERFORMANCE OF TWO SUBGROUPS IN A SERIES OF DAILY REVERSALS AFTER 12 CRITERION-REVERSALS IN THE CONTROL KEY-DISCRIMINATION

C, blue-green discrimination with magazine-light the color of the positive key; *W*, blue-green discrimination with white magazine-light; *R*, blue-green discrimination with magazine-light the color of the negative key; *RY*, red-yellow discrimination with magazine-light the color of the positive key; *BY*, blue-yellow discrimination with magazine-light the color of the positive key.

magazine produced a much greater than chance-level. In the next 20 sessions, however, performance improved steadily, and, when they were shifted back to the control condition in the thirty-eighth session, the C-1 animals showed another disruption of performance. With the white magazine in the last three sessions, the level of error was not much better than chance.

DISCUSSION

Our principal conclusion from these results is that the ingestive technique offers no special advantage over key-pecking insofar as the rate of discriminative learning is concerned. While the performance of the magazine-animals was much superior to that of the standard key-animals (trained with a white magazine-light), it was not better than that of the control key-animals (trained with a magazine-light of the same color as

the positive key). The similarity of the magazine- and control-performances suggests that pecking a key is as effective an indicator-response as choosing a magazine, and that nothing is gained from a close proximity of the locus of response to the locus of reward which the ingestive technique affords. The superiority of the magazine- and control-techniques over the standard key-technique probably is due to the contiguity of positive cue and reward which both the former provide and the latter does not. While the procedures employed for the magazine- and control-animals differed in certain other respects from those employed for the standard key-animals, the deterioration in the performance of the control animals when they were shifted to the standard condition demonstrates the importance of cue-reward contiguity.

A technique reminiscent of the present control technique has been used effectively in work with rats. Grice found that the learning of a black-white discrimination would be impaired if an interval of delay was interposed between the animal's choice and its admission to a gray goal-box (if correct) or an empty gray end-box (if incorrect); the detrimental effect of delay was reduced, however, if the goal-box was the color of the positive cue and the empty end-box the color of the negative cue.[6] In his interpretation of this finding, Grice relied on the notion of secondary reinforcement, but the results of a subsequent experiment showed that the matter is not quite so simple. Bauer and Lawrence found better performance in a group trained with a goal-box the color of the *negative* cue and an empty end-box the color of the *positive* cue than in a group trained with identical boxes.[7] Also to be considered in this connection is the fact that the *C*-1 animals of the present experiment, having adjusted to the reverse-magazine condition, made many errors when shifted back to the positive-magazine condition, and still were unable to perform at a better than chance level under the white-magazine condition. The colored goal-box or magazine apparently serves to define the correct stimulus in a more general way than contemplated by Grice.

A possibility which suggests itself is that the magazine- and control-animals could improve their performance without learning to reverse, but simply by learning to follow the magazine-color. To account in these terms for the superiority of the magazine- and control-animals, it is necessary, of course, to assume that delayed matching is more readily accomplished than reversal learning. The color-specificity of the control performance, which is indicated by the rise in the error-score of Group *C*-2 when it was shifted from blue-green to red-yellow, certainly is reminiscent of the color-specificity which has been reported for matching by Cumming and Berry-

[6] G. R. Grice, The relation of secondary reinforcement to delayed reward in visual discrimination learning, *J. exp. Psychol.*, 38, 1948, 1-16.

[7] F. J. Bauer and D. H. Lawrence, Influence of similarity of choice-point and goal cues on discrimination learning, *J. comp. physiol. Psychol.*, 46, 1953, 241-248.

man.[8] That the poor red-yellow performance of the C-2 animals was not due simply to the lesser discriminability of the two new colors is indicated by the fact that performance did not deteriorate until the *second* red-yellow session (the first red-yellow reversal). It is conceivable, however that progressive improvement in reversal-learning is as color-specific in the pigeon as is matching. Evidence against the matching-interpretation is provided by the good performance of the C-2 animals on the first day following each of the two shifts from the control to the standard (white-magazine) condition. Why the deterioration in performance should have been delayed in each case is a difficult question. It is clear that much remains to be learned about the role of cue-reward contiguity in discriminative learning.

Summary

An ingestive technique, which offers the pigeon a choice between two differentially colored food-magazines, was compared with a standard key-pecking technique and found to give much more rapid discriminative learning, but a control technique, which differs from the standard key-pecking technique only in that the magazine-light is the same color as the positive key, was found to be as efficient as the ingestive technique. It was concluded that the ingestive indicator-response has no special advantage over key-pecking provided that cue-reward contiguity is equated. Some further results on the role of a colored magazine in the key-color discrimination also were considered.

[8] W. W. Cumming and Robert Berryman, Some data on matching behavior in the pigeon, *J. exp. Anal. Behav.*, 4, 1961, 281-284.

BINAURAL SUMMATION OF THERMAL NOISES OF EQUAL AND UNEQUAL POWER IN EACH EAR

By R. J. Irwin, University of Auckland, New Zealand

It is well established that a sound heard with two ears is judged louder than the same sound heard with one ear.[1] A similar effect has been found at threshold: A liminal stimulus for binaural listening is usually lower than that for either ear alone. Both these effects are called binaural summation, but the experiment reported here is concernd only with summation at levels above threshold. In particular, it investigates what combinations of thermal noises of equal and unequal power are equivalent in loudness to a comparative monaural noise. The difference between the total power of the binaural and of the monaural sounds served as a measure of the superiority of binaural over monaural listening.

The words 'total power' of the binaural stimulus are used here to mean the sum of the powers presented to each ear. It is important to make clear how the power in each ear is specified. If the two ears of a listener are not equally sensitive (and they rarely are), stimuli of equal sound-pressure level (*SPL*) in each ear will not be of equal loudness, unless the growth of loudness with *SPL* is also different from ear to ear. For this reason, the *SPL* in each ear is not necessarily the appropriate property of the stimulus for binaural summation. For example, if a listener were very deaf in one ear, one would not expect to find binaural summation when comparing the loudness of a sound in the good ear with its loudness when the same total power was presented to both ears. Indeed, binaural subtraction would be expected, since that portion of the monaural power delivered to the deaf ear would be lost, insofar as its sensory effect was concerned. The same argument holds in less extreme cases, so that the pertinent specification of stimulus-level is in decibels above the threshold of each ear; that is, in decibels of sensation-level (*SL*). 'Total power' means, therefore, the sum of the powers in each ear, when they are stated in decibels above the threshold of each ear.

An alternative method, which Fletcher and Munson adopted,[2] is to assume that if a sufficient number of listeners are used, interaural differences in sensitivity will on the average cancel out, and sounds of equal *SPL* will on the average be of equal loudness.

---

* Received for publication October 4, 1963. My thanks are due Messrs. E. E. Dent and A. Ross for their help in designing and maintaining the equipment.

[1] I. J. Hirsh, Binaural summation—A century of investigation, *Psychol. Bull.*, 45, 1948, 193-206.

[2] H. F. Fletcher and W. A. Munson, Loudness, its definition, measurement, and calculation, *J. acous. Soc. Amer.*, 5, 1933, 82-108.

Most previous studies of binaural summation have measured the effect when the power of the binaural sound was the same in each ear. An exception was an experiment by Hughes on binaural summation at threshold.[3] He measured the threshold of each ear separately, then presented a stimulus to one ear that was 1, 3, or 5 db. below its threshold, and redetermined the threshold in the contralateral ear. In effect, he was now measuring a binaural threshold. The liminal stimulus in the contralateral ear proved to be such that the total power in the two ears remained the same, regardless of how it was divided between them. His results indicated that at levels near threshold a constant effect was produced by a constant total power.

At suprathreshold levels, however, a constant total power does not necessarily produce a constant effect.[4] Reynolds and Stevens reported,[5] for example, that a binaural sound of 90 db *SPL* in each ear was as loud as a monaural one of 98 db. *SPL*—not 93 db. *SPL* as constant total power would dictate.

There is a region, however, between the two cases of all the power going to one ear and equal powers going to each, for which the resultant loudness is unknown. The object of the experiment reported here was to map this region at four different levels.

Method and Procedure

Apparatus. A schematic diagram of the basic apparatus is shown in Fig. 1. The noise generator, N, was of the gas tube type. Its output could be switched on and

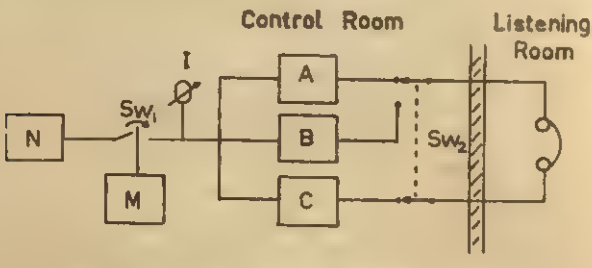

Fig. 1. Schematic Diagram of the Apparatus

off by the relay switch Sw_1 which was operated by a cam driven by the motor, M. The output of the generator was monitored by the ammeter, I. The switch Sw_2 allowed the noise to be delivered either to both phones through Attenuators A and C, or to the left phone alone, through Attenuator B.

The motor rotated once every 12 sec. during which time two signals were presented for about 0.5 sec. each, separated from each other by about 1.5 sec. The

[3] J. W. Hughes, The monaural threshold: Effect of a subliminal contralateral stimulus, *Proc. roy. Soc.*, 124 B, 1938, 406-420.

[4] Fletcher and Munson, *op. cit.*, 82-108.

[5] G. S. Reynolds and S. S. Stevens, Binaural summation of loudness, *J. acous. Soc. Amer.*, 32, 1960, 1337-1344.

remaining 9.5 sec. of the 12-sec. cycle was silent. During the 1.5 sec. which separated the two signals, Sw_2 could be switched (manually) that one of the signals could go to the left ear and the other to both ears.

A lamp in the listening room was turned on and off by Sw_1 so that a light came on together with the sound. In this way the presence of the auditory stimulus was unambiguously indicated by the light.

Two buttons in the listening room each illuminated one of a pair of lights in the experimental room, together with one of a similar pair in the listening room. The subject (S) communicated his judgment to the experimenter (E) by pressing these buttons.

Subjects. Four men, second year students in psychology, volunteered to serve as *Ss*. Their ages were 19, 26, 22 and 31 yr., and they are hereafter referred to as S_1 to S_4, respectively. Their pure tone sensitivity was measured at 10 frequencies with an Amplivox 82 Audiometer. S_3 had a loss of 25 db. for a 4000-~ tone in his right ear, and a loss of 30 db. for a 6000-~ tone in the same ear. Apart from this, none of the *Ss* had a loss of more than 20 db. in either ear at any of the frequencies measured.

Procedure. Three thresholds were determined for each S by the method of limits. The first threshold was for the noise delivered to the left ear through Attenuator A; a second threshold for the noise in the left ear delivered through Attenuator B; and a third for the noise in the right ear delivered through Attenuator C. The power delivered to the left phone through Attenuator B was then raised to 10-, 30-, 50-, and 70-db. *SL* in an order that was random and different for each S. Comparisons were made for each of these stimuli with one delivered to the left ear through Attenuator A, and then with one to the right ear through Attenuator C. The settings of Attenuators A and C which were separately equal to that of Attenuator B were thereby found.

In determining the setting for equal loudness, S reported which of the two stimuli was the louder, the first or the second, or whether they were the same. E adjusted the appropriate attenuator, usually in steps of 1 db., until S judged the two noises to be of equal loudness. The psychophysical method was thus the method of adjustment, in which the experimenter, rather than S, adjusted the stimulus.[6] The constant stimulus through Attenuator B was presented first half the time, and second half the time. The variable stimulus; namely, the stimulus delivered through Attenuator A or C, was started as often above the constant stimulus as below it.

S next compared a binaural stimulus with a monaural one. The loudness of the noise in the left phone delivered through Attenuator B at one of the four *SLs* was compared with the binaural noise presented through Attenuators A and C when each of these was set at a level previously judged equal to B. Attenuator B was adjusted until the monaural sound in the left ear was equal in loudness to the binaural one. The difference in decibels between the original setting of B and the setting for equality with the binaural noise gave a measure of the summation of noises of equal power in each ear.

This setting of B was then used as a comparison-level to determine what other

[6] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*. (2nd ed.), 1954, 200.

combinations of A and C produced the same effect. The power delivered to the left phone, for example, was reduced by a predetermined number of decibels, and *S* judged what increment in the right phone was necessary to produce a constant effect; namely, an effect equivalent in loudness to the comparison stimulus in the left ear. The specific amount by which the power in one phone was reduced was so chosen that there were many combinations near the point of equal power in each ear, and a few very unequal combinations. For example, when the comparison-stimulus was set equal to a binaural sound of 50-db. *SL* in each ear, *S* was given the following series of powers in the left ear: 20-, 30-, 40-, 42-, 44-, 46-, and 48-db. *SL*. Each of these was presented in turn with 50-db. *SL* in the right ear. The power in the right ear was then adjusted until the binaural combination was judged equal to the monaural comparison-sound. The same series of *SLs* was also presented to the right ear, while 50-db. *SL* was given to the left. The power in the left ear was then adjusted until the binaural pair was judged equal to the comparison-sound. The actual order in which any combination was presented to *S* was random and different for each *S*, and for each comparison-level, except that all the reductions from a given *SL* were made in the left ear before they were made in the right.

The judgments of each *S* were made on seven separate occasions, each lasting about an hour. All judgments for combinations of unequal power at one of the comparison-levels were made in a single session. The earphones were reversed on S_2 and S_4 with respect to their position on S_1 and S_3.

Results

The average summation-effect for sounds of equal power at the four *SLs* is shown in Fig. 2. The ordinate shows the *SL* of the binaural sound in each ear. The extent to which the empirical points fall below the diagonal shows the superiority of binaural over monaural listening. This superiority tends to increase with *SL*, a result which agrees with that of Fletcher and Munson and of Reynolds and Stevens,[7] but not that of Caussé and Chavasse, who found the decibel-difference between equally loud binaural and monaural sounds remained constant after about 35-db. *SL*.[8] An extrapolation of the empirical data down to the binaural threshold would intersect the monaural *SL* at about 3 db, which is a typical finding for the decibel-difference at threshold, provided the sensitivity of each ear is equated.[9]

The same data are shown in Fig. 3, to illustrate what proportion of the monaural power was judged equally loud when heard binaurally. At high levels, the total power of the binaural noise was only about one-

---

[7] Fletcher and Munson, *op. cit.*, 82-108; Reynolds and Stevens, *op. cit.*, 1337-1344.
[8] R. Caussé and P. Chavasse, Différence entre l'écoute binauriculaire et monauriculaire pour la perception des intensités supreliminaire, *C. R. Soc. Biol.*, Paris, 136, 1942, 405.
[9] Hughes, *op. cit.*, 406-420.

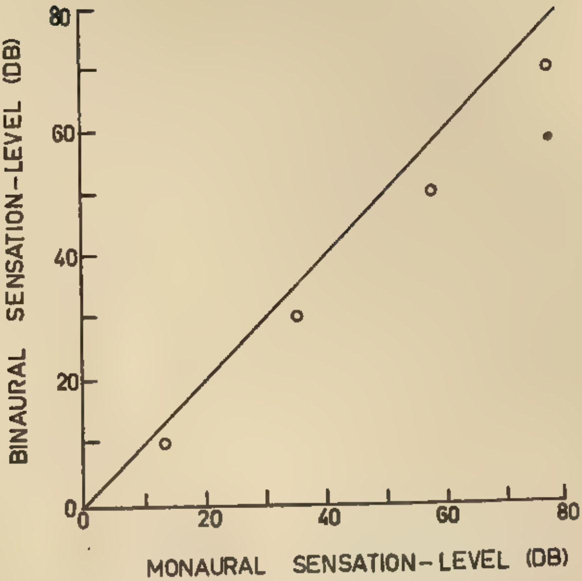

FIG. 2. EQUALLY LOUD BINAURAL AND MONAURAL NOISES
Each component of the binaural stimulus is the same level above threshold. The straight line is drawn through the diagonal at 45°.

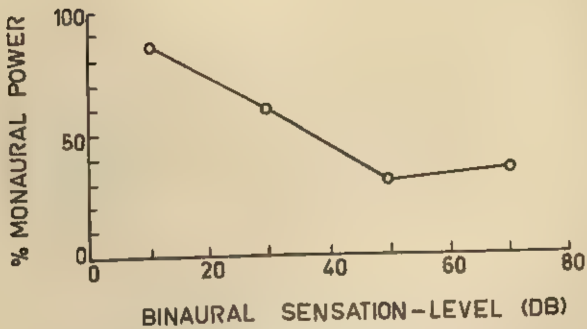

FIG. 3. TOTAL POWER OF THE BINAURAL STIMULUS EXPRESSED AS A PERCENTAGE OF THE EQUALLY LOUD MONAURAL STIMULUS

third of that of the equally loud monaural noise. At 10-db. *SL* in each ear it was about four-fifths of the monaural noise.

The summation of noises of unequal power for S_1 is shown in Fig. 4. The results of a single *S* are illustrated because the averaged results obscure some of the trends. Each experimental curve in this figure is a con-

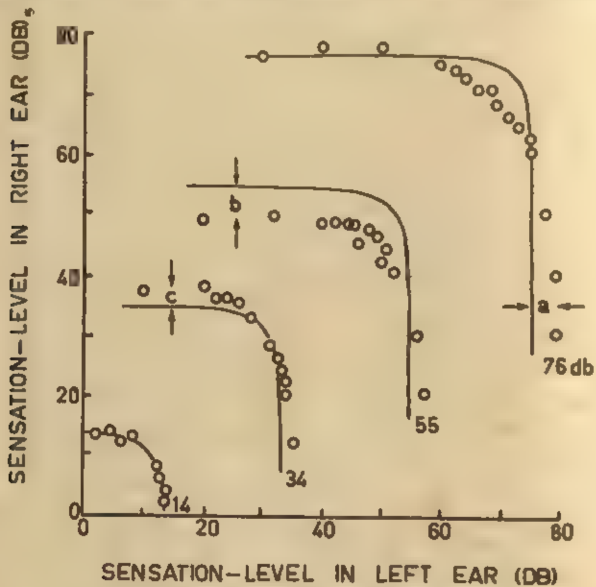

Fig. 4. Contours of Constant Loudness
The data are from one subject. The parameter is the *SL* of the comparison-stimulus. The solid lines show what combination in each ear maintains a total power equal to that of the comparison-stimulus.

tour of constant loudness. The parameter is the *SL* of the monaural comparison sound judged equal to a binaural noise of 10-, 30-, 50-, and 70-db. *SL* in each ear. The experimental points represent combinations of binaural sounds that were judged equally loud to the comparison-sound. The solid curves are contours of constant total power. That is to say, they show what combinations of power in each ear maintain a power equivalent to that of the monaural stimulus alone.

The experimental points for the 76-db. comparison-level depart most from the curve of constant power when the binaural stimuli are about equal; but when the level in the left ear is reduced to 60-db. *SL*, the total power needed for equality approaches that of the monaural power. Similarly, as the power in the right ear is reduced to 60-db. *SL*, the total binaural power approaches the monaural power.

The distance, marked 'a' on Fig. 4, between the monaural power and

the total binaural power at very unequal combinations, needs consideration. The results indicate that a power of 30-db. *SL* in the right ear and one of 80-db. *SL* in the left ear were judged the same loudness as one of 76-db. *SL* in the left. In this case, the contribution of the power in the right ear to the total binaural power is, of course, negligible (the power in the right ear is 10^{-5} times less than that in the left), so that the results in effect indicate that a noise of 80-db. *SL* in the left ear equals one of 76-db. *SL* in the left. The two noises were delivered through different attenuators, and the setting of each attenuator for equality at 70-db. *SL* of the comparison-noise had been determined in a previous experimental session. The fact that they differ by 4 db. when the comparison-level is 76-db. *SL* implies that judgments of equality differ systematically from one occasion to another. Apart from *S*'s inherent variability, this might be partly the result of a different coupling of the receiver to the ear from one occasion to another. The distances marked 'b' and 'c' at the 55- and 34-db. comparison-levels may be given a similar interpretation.

The effect of these systematic shifts is cancelled out to some extent when the judgments of all four *S*s are averaged; the general effect is thereby made clearer. The averaged data are presented in Fig. 5. The parameter is the average *SL* of the comparison-noise. (The two extreme points of the 35.25-db. contour are based on only 3 *S*s, as those points were inadvertently not found for S_3.) Again, the greatest advantage of binaural over monaural listening is found for approximately equal powers at high levels. As the powers are made more and more unequal, so the advantage of binaural over monaural listening declines.

A sound reduced 10 db. in either ear at the two higher levels continues to contribute to the total effect, in the sense that the binaural power is less than that of the equally loud monaural power. At the 35.25-db. level, however, the contribution of a sound reduced 10 db. largely disappears. At the 13.75-db. level, the summation-effect is at best small, and the points follow a course parallel to the line of constant monaural power. A systematic lessening of the binaural advantage with increasing inequality of each binaural component cannot be observed at the 13.75-db. level.

Discussion

One simple interpretation of the findings rests on the concept of critical bandwidth which Fletcher inferred from the masking of pure tones by white noise, and which Zwicker, Flottorp, and Stevens measured directly.[10]

---

[10] H. F. Fletcher, Auditory patterns, *Rev. mod. Phys.*, 12, 1940, 47-65; E. Zwicker, G. Flottorp, and S. S. Stevens, Critical band width in loudness summation, *J. acous. Soc. Amer.*, 29, 1957, 548-557.

The critical band defines a frequency-region within which loudness is independent of bandwidth. The loudness of noises whose spectra are narrower than a critical band has been shown to depend only on their power; but the loudness of noises whose spectra are wider than a critical band depends both on their power and on their bandwidth. The width of a

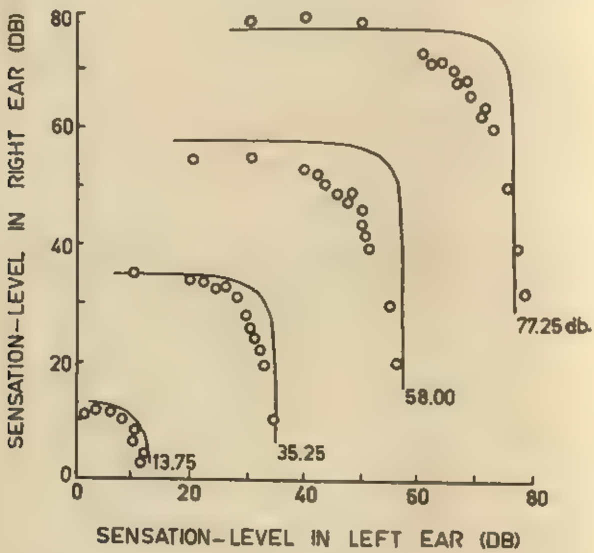

FIG. 5. CONTOURS OF CONSTANT LOUDNESS
The data are the average of four *Ss*. The parameter is the average *SL* of the comparison-stimulus. The solid lines show what combinations in each ear maintain a total power equal to that of the comparison-stimulus.

critical band varies with its center frequency and it is thought to reflect some fundamental property of the basilar membrane.

An interpretation of binaural summation can be made on the supposition that stimulating two ears amounts to an increase in the bandwidth of the noise, in the sense that it is now spread over two basilar membrances. If stimulating two different regions of the same membrane is more effective than stimulating one region, so stimulating two membranes might be expected to be more effective than stimulating one, for the same reason. Because bandwidth is in effect increased in binaural stimulation, power must be decreased if loudness is to remain unchanged.

The gradual convergence of the total binaural power and the power of the comparison-noise can be interpreted as the declining effect of added bandwidth as its power is reduced. If insufficient power is added, the in-

crement in loudness falls below the *JND*, and the contour of constant loudness settles on the monaural level. One would expect this *JND* to be smaller than that of an increment in power that did not also involve an increment in bandwidth. The *JND* for the loudness of noise, when power alone is increased, is about 10% at high levels,[11] so that if added bandwidth were a factor in binaural summation, a binaural stimulus with an interaural difference of 10 db. should be more than noticeably different from the louder one alone. In the two upper contours of the averaged data, an interaural difference of 10 db. requires less power than an equally loud monaural noise, and this suggests that a 10% increment in power to the other ear is above the *JND*, as an explanation based on added bandwidth predicts. The point needs explicit checking, however.

Summary

The relation between binaural and monaural loudness was studied by determining what binaural combinations of noises of equal and unequal power—measured in decibels above threshold—were judged equivalent in loudness to a comparative monaural stimulus. Four *S*s, men, made judgments of equality at each of four levels of the comparison-noises—levels equivalent to binaural noises of 10-, 30-, 50-, and 70-db. *SL*. Contours of constant loudness, for each of the four levels of the comparison-stimulus, indicated that the advantage of binaural over monaural listening was greater at high levels than at low. The advantage was also greater, at any comparison-level, when the binaural noise had about the same power in each ear. At high levels with equal binaural stimuli, the total binaural power was about one-third of that of the equally loud monaural stimulus, and at low levels it was about four-fifths. As the difference between the two components of the binaural stimulus was increased, the powers of equally loud binaural and monaural noises gradually converged.

The superiority of binaural over monaural listening was interpreted to be similar in nature to the increment in loudness that is observed when the bandwidth of a monaural noise is increased beyond a critical value.

[11] G. A. Miller. Sensitivity to changes in the intensity of white noise and its relation to masking and loudness, *J. acous. Soc. Amer.*, 19, 1947, 609-619.

INTERMANUAL EFFECTS OF ANCHORS ON ZONES OF MAXIMAL SENSITIVITY IN WEIGHT-DISCRIMINATION

*By DOROTHY DINNERSTEIN, Rutgers University

Suppose that *S*, using his *right* hand, compares by the method of constant stimuli an 80-gm. standard weight with a variable series ranging from 70-90 gm. Will these discriminations be affected if at each lift he simultaneously lifts with his *left* hand a non-judged 'anchor' weight of 320 gm. or of 20 gm.?

In view of what is known in other connections about the dependence of local perceptual events on the properties of the surrounding stimulus-field, one might be surprised to find discrimination unaffected. It may be, however, that such an influence cannot extend from one hand to the other; and perhaps for weight-difference perceptual items not directly involved in the comparison are irrelevant. Non-judged anchors do, to be sure, strikingly affect weight-judgments when the method of single stimuli is used.[1] This fact, however, can—as Stevens has recently emphasized—be interpreted as reflecting not a change in "sensory excitability" but a purely semantic shift.[2] *S* must distribute the five or seven verbal labels at his disposal equitably over the entire range of magnitudes on which the experiment focuses his attention; when this range is changed by a lighter or heavier anchor, it may be that a given weight can move into a different fifth or seventh, acquiring a new label, or that two different weights can crowd into the same fifth or seventh, coming to share the same label, while *Ss*' actual perception of magnitudes and awareness of differences between them remains unchanged. If the single-stimuli anchor-effects are in fact due solely to such semantic adjustments and not to perceptual shifts, then in the present experiment, where *S*'s task is to make discriminations rather than to assign labels, the anchor should prove ineffectual.[3] If, however, it is true, as Helson suggests,[4] that in

* Received for publication December 4, 1963.

[1] Harry Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this JOURNAL, 60, 1947, 1-29.

[2] S. S. Stevens, Adaptation-level vs. the relativity of judgment, this JOURNAL, 71, 1958, 633-646.

[3] It has been demonstrated by J. P. Guilford and Dorothy G. Park (The effect of interpolated weights upon comparative judgments, this JOURNAL, 43, 1931, 589-599) that light and heavy anchors interpolated between standard and variable with the method of constant stimuli, lower and raise, respectively, the *PSE*. These

weight-judgment as in color-vision S's discriminative capacities are somehow flexibly focused on the center of the currently-attended-to range of stimuli, a substantial enough shift in the location of the center of this range should result in a shift in the zone within which S makes his most sensitive judgments.

Method

Using the right hand with the elbow resting on a table, S, seated and blindfolded, lifted first a standard weight of 80 gm., and then one of five comparison-weights—70, 75, 80, 85, and 90 gm.—and stated whether the second seemed "lighter" or "heavier" than the first. If the two seemed equal, S guessed, saying "guess lighter" or "guess heavier."[5]

The weight, a plastic pillbox filled with cotton and shot, was standing directly under S's hand when E, who used a turntable to facilitate the presentation, said "Now" or "OK." S then lifted it several inches off the table with a relaxed wrist, and put it down again at a signal from E. A period of 2 sec. was allowed for this act, followed by a 3-sec. interval between standard and comparison-weights and a 6-sec. interval between trials, during which S stated his judgment and E recorded it and prepared the weights for the following trial.

S made 2,000 such judgments, four blocks of 500 each, since each of the five comparison-weights was thus presented for paired successive comparison with the standard 100 times (the series of five occurring in 100 pre-arranged random orders) under each of four conditions: In the "heavy anchor" condition, each time S lifted the standard or a comparison-weight with his right hand, he simultaneously lifted a 320-gm. weight with his left. In the 'light anchor' condition, the weight in the left hand was 20 gm. In the 'equal anchor' condition, it was 80 gm., matching the standard. In the 'no anchor' condition, the left hand was idle. Two such blocks of 500 trials were carried out in a single session and the other two in a second session a day or two later.

The Ss were four college graduates in their twenties, unaware of the purpose of the experiment and inexperienced in this type of task. The order in which the anchor-conditions occurred was varied as follows: (1) none, equal; light, heavy. (2) equal, none; heavy, light. (3) heavy, light; equal, none. (4) light, heavy; none, equal.

Results

In Table I, results are presented in terms of accuracy of discrimination, *i.e.* numbers of wrong (vs. correct) judgments. 'Equal' judgments having

---

investigators were puzzled to discover differing effects at the upper vs. lower ends of the comparison-series—a fact consistent with the results to be presented here and with Helson's theory. Since their anchors were temporally interpolated between the two weights, S was comparing, however, their central finding can be interpreted as bearing mainly on the problem of time-error.

[4] Helson, Adaptation-level as a basis for a quantitative theory of frames of reference, *Psychol. Rev.*, 55, 1948, 297-313.

[5] Data for this experiment were collected by Mrs. Carol Gordon under a grant from the Rutgers University Research Council in the fall of 1961. I am also indebted to Dr. Mavis Hetherington, who acted as E in an exploratory experiment, for her interest in and help with this study.

been forbidden, this analysis for the moment excludes the 80-gm. variable.

In considering Table I, it should be noted first that the equal (80-gm.) anchor, as compared with no anchor, consistently reduces accuracy of discrimination, making for a higher mean number of errors for every variable weight on the table. (Of the 16 individual pairs of error-totals involved, for four *S*s and four variable weights, 13 went in this direction, with one tie and two scattered exceptions.) Since this anchor does not alter the range of weights attended to by *S*, its effect is presumably attributable to the presence *per se* of a non-judged weight in the other hand. Hence it provides a baseline for comparison with the anchors which do alter the range. Note, then, that as compared with the equal anchor, the light and heavy ones make for finer and coarser discrimination, respectively, at the

TABLE I

MEAN NUMBERS OF ERRORS IN DISCRIMINATION UNDER FOUR ANCHOR-CONDITIONS

Variable weight (gm.)	Anchor (gm.) none	80	20	320
70	10.7	12	6.5	17.7
75	21	23.5	20.2	26.7
85	18.5	22	21.7	19.7
90	4.5	5.7	6.2	4.7

light end of the variable series: the 70- and 75-gm. variables were *more* often incorrectly judged with the heavy than with the equal anchor (seven of eight individual pairs of error-totals going in this direction), and *less* often incorrectly judged with the light than with the equal anchor (all eight individual pairs of totals going in this direction). A corresponding, though weaker and less reliable, trend occurs at the heavy end of the series, with the heavy anchor slightly improving discrimination for both 85- and 90-gm. variables and the light anchor slightly coarsening it for the 90-gm. variable. These findings, then, support Helson's view that in weight-judgment, as in perception of brightness and hue, shifts in sensory excitability follow shifts in the location of the center of the range of values to which *S* is attending.

Table II summarizes these facts in another way. It presents ratios comparing accuracy of discrimination of differences when the variable is lighter, vs. heavier, than the standard. Thus the first row shows that with no anchor, and to a slightly reduced extent with the equal anchor, discrimination is more accurate when the variable is heavier, *i.e.* more wrong judgments are made below than above the standard weight: error ratios for both 5- and 10-gm. differences are larger than one. Ratios in this

direction are to be expected with our procedure, since the time-error in successive comparison when the standard comes 3 sec. before the variable should work to increase the apparent relative weight of the latter, thus enhancing the difference when it is objectively heavier and subtracting from it when it is objectively lighter. Note, however, that the size of these ratios—*i.e.* the degree of superiority in discriminability of weights at the heavy end of the series—is sharply *increased* with the *heavy* anchor, and *reduced* with the *light* anchor.

One might speculate from these figures that the anchors have differing effects on the time-error *per se*, the heavy anchor making for a larger time-error, hence more "heavier" judgments. Such an interpretation, which would in any case be strangely at odds with what is known about

TABLE II

Ratios of Errors in Discrimination when Variable Is Lighter or Heavier than the 80-gm. Standard under Four Anchor Conditions

Difference from standard	Anchor (gm.)			
	none	80	20	320
5 gm.: 75/85 gm. variables	1.13	1.07	0.93	1.36
10 gm.: 70/90 gm. variables	2.38	2.11	1.05	3.77
5 and 10 gm.: 70- and 75-/85- and 90-gm. variables	1.75	1.59	0.99	2.56

the effects of varying background-weights on the direction of the time error,[6] is ruled out by results for the 80-gm. variable, which have not so far been considered. Fig. 1, comparing the two extreme anchor-conditions in terms of "lighter" (vs. "heavier") rather than wrong (vs. correct) judgments, reveals the interesting fact that while for every comparison-object differing in weight from the standard, the mean number of "lighter" judgments is greater with the 20- than with the 320-gm. anchor, when the comparison object *equals* the standard, the difference goes in the opposite direction. Why should the anchors exert this exceptional effect on judgment with the 80-gm. variable, when they seem to act otherwise with variables both heavier and lighter than the standard?

Two possible resolutions of this paradox suggest themselves. On the one hand, results for the 80-gm. variable may be taken as an index of hypothetical differences among anchor-conditions as to time-error, in which case the light and heavy anchors—as they do when interpolated—enhance and reduce, respectively, the negative time-error. If this is true, then time-error effects, far from *explaining* the differences for the other four variables shown in Tables I and II and in Fig. 1, actually work to *blunt* them, the actual contrasting effects of anchors on discrimi-

[6] Guilford and Park, *op. cit.*, this Journal, 43, 1931, 589-599.

nation at the near vs. far ends of the variable series being in fact more marked than these data show them to be. On the other hand, it may be that these anchors, since they were not interpolated but lifted simultaneously with both standard and variable, with an empty interval intervening, did not affect the time-error at all. If this is assumed to be possible, a key to the puzzle of Fig. 1 emerges: Table III shows, first, that with no anchor there is—as would be expected with this procedure—a slight preponderance of "heavier" judgments for the 80-gm. variable.

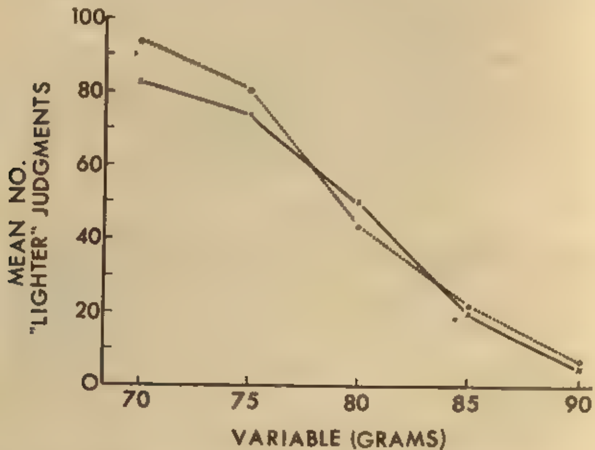

FIG. 1. MEAN PERCENTAGE OF "LIGHTER" JUDGMENTS WITH 20-GM. (BROKEN LINE) AND 320-GM. (UNBROKEN LINE) ANCHOR

TABLE III

MEAN NUMBERS OF "LIGHTER" JUDGMENTS (ERRORS IN DISCRIMINATION?) UNDER FOUR ANCHOR-CONDITIONS WHEN VARIABLE = 80 GRAMS = STANDARD

	Anchor (gm.)			
	none	80	20	320
Variable = 80 grams	41.0	47.8	43.2	49.7

This is an expression of the normal negative time-error; *i.e.* through assimilation to the empty intervening interval the standard as preserved in recent memory is phenomenally slightly lighter than the variable as currently perceived. The difference is small, and may be assumed to become clear only under optimal perceptual conditions, but it is a genuine difference; hence "lighter" judgments may be viewed as, in a sense, less correct than "heavier" judgments. If this is regarded tentatively as true regardless of the weight in *S*'s left hand, then anchor-conditions can be compared for effects on discrimination with the 80-gm. variable as with the others. Note first that the mean number of 'errors' is larger with the equal anchor than with none (all four individual pairs of 'error'-scores were in this direction). Thus here, as with every other variable weight, this anchor re-

duces accuracy of discrimination. Taking the equal anchor, again, as an index of the influence of the presence *per se* of a non-judged weight in the left hand, and gauging against it the effects of anchors which in addition alter the range of attended-to weights, one then notes that with the 80-gm. variable, as with the 70- and 75-gm. variables, the light anchor makes for fewer 'wrong' judgments and the heavy anchor for more (again, for both comparisons, with no individual exceptions). Thus data for the 80-gm. variable confirm the finding that the light and heavy anchors increase and reduce, respectively, accuracy of discrimination at what

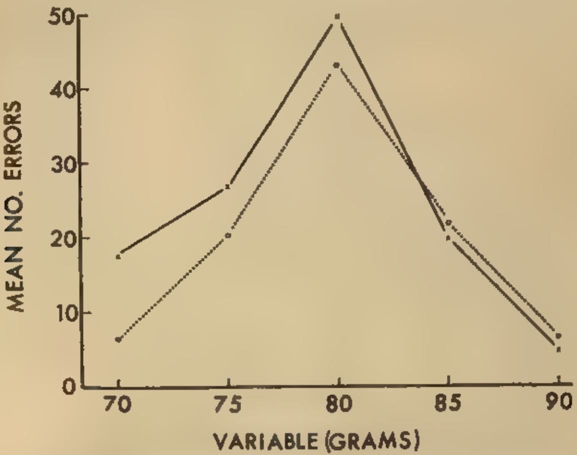

FIG. 2. MEAN PERCENTAGE OF ERRORS IN DISCRIMINATION WITH 20-GM. (BROKEN LINE) AND 320-GM. (UNBROKEN LINE) ANCHORS

(At the 80-gm. variable, the assumption is made—see text—that "lighter" judgments represent errors.)

is for the former the near, and for the latter the far, end of the variable series, and the paradox of Fig. 1 is resolved. See Fig. 2.

Further work is needed to decide between these alternate explanations. Whichever proves correct, it is clear that results with the 80-gm. variables rule out an explanation of the main findings in terms of time-error differences between anchor-conditions, and strengthen the conclusion that these findings reflect consistently contrasting effects of the anchors on discriminability.

In the comparisons presented thus far, guessing judgments were not separately analyzed, "guess heavier" and "guess lighter" judgments being included in the "heavier" and "lighter" totals respectively.[7] Additional

[7] Only three of the four *Ss* gave guessing judgments. Those given were, of course, correct strikingly more often than chance: of 48 sets of guessing judg-

confirmation for the main findings already discussed appears when guessing judgments are considered in their own right, the percentage of all correct judgments which *S*s preceded by the qualification "guess" being calculated for the 70-, 75-, 85-, and 90-gm. variable weights under each anchor-condition. When these values are examined for the light- and heavy-anchor conditions, each anchor as compared with the other proves to make for a smaller percentage of "guessing" judgments among correct judgments at the near end of the variable series, and a larger percentage at the far end. Thus, not only is *accuracy* of discrimination relatively better at the near end, but correct discriminations when made are made with greater *certainty*.[8]

Exploratory experiments indicate that patterns of effects different from those reported for the present anchors would probably appear with more, or less, extremely light and heavy anchors.[9] Further study is needed to make it clear how an anchor's effects, with the present method, can be expected to be distributed within the altered range which it creates. The only conclusion offered here, then, is the general one that a shift in the location of the range of attended-to weights of which a given series of weight-differences is part results in a shift in the zone of maximal discriminability within the series.

DISCUSSION

This study, like others which have appeared recently,[10] suggests that the effects of anchors on judgment found with the method of single stimuli

---

ments—three *S*s four variable weights unequal to the standard four anchor-conditions—43 showed a preponderance of correct over incorrect judgments, with four ties and two scattered exceptions.

[8] Besides confirming this central finding of the present study, analysis of guessing judgments also confirms the report of Guilford and Park (*op. cit.*, 589-599) that both heavy and light anchors, as compared with an anchor equal to the standard, increase the over-all difficulty of the discriminative task. This finding—*i.e.* that apart from the location of its center the sheer breadth of the range of attended-to weights affects subjective certainty—will be presented more fully when the problem on which it bears has been explored further with the present anchor-technique.

[9] This would correspond to findings with the method of single stimuli. (Harry Helson and C. Nash, Anchor, contrast, and paradoxical distance effects, *J. exp. Psychol.*, 59, 1960, 113-121).

[10] O. J. Harvey and D. T. Campbell, Judgments of weight as affected by adaptation range, adaptation duration, magnitude of unlabelled anchor, and judgmental language, *J. exp. Psychol.*, 65, 1963, 12-21. William Bevan, R. A. Maier, and Harry Helson, The influence of context upon the perception of number, this JOURNAL, 76, 1963, 464-469. Both these studies demonstrate effects of the phenomenal surround on direct judgments of amount: in the former study, weights are estimated in ounces and in the latter batches of beans assessed as to numerosity.

must, as Helson thought, reflect some genuine perceptual shifts.[11] It by no means follows that the semantic shifts to which Stevens refers[12] were not also involved,[13] or that the distinction between changed modulus of judgment and changed sensory excitability is not important. What motivated the present study was, rather, an interest in the question of whether these two distinct kinds of changes—taking place by way of different mechanisms, at different levels of organization of the nervous system, but functionally similar in that both involve a shifted distribution of cognitive-perceptual effort in response to a shifted stimulus-range—can actually occur in so sharply separate a way as Stevens seems to imply.

Köhler has recently discussed and experimentally illustrated the fact that the intensity of a perceptual process varies with attention.[14] This means that the distribution of energy to any given local phenomenon depends on the organization of the larger experiential field of which it is part. Many recent studies could be cited as examples: a figural after-effect can be shaped by the memory-images with which *S* groups the unseen object he holds in his fingers;[15] sensitivity to a difference between two visual figures with respect to one perceptual characteristic (microstructure) decreases with an increase in the sharpness of a difference between them with respect to another characteristic (size),[16] and so on. From this point of view it seemed unlikely, in the case of the single-stimuli anchor-experiments referred to by Helson in support of his *AL*-concept, that the judgmental effect could fail to be accompanied by a perceptual one; for a shift in the stimulus-range large enough to change the modulus of judgment should also cause a redistribution of attention sufficient to affect discriminability. This experiment was undertaken to explore what seemed the strong probability that if anchors can change the phenomenal situation enough to affect the names *S* gives to his weight-experiences, the change also must affect the experiences themselves.

---

[11] Helson, *op. cit.*, this JOURNAL, 60, 1947, 1-29. In support of this view of his data, Helson cites an important fact (p. 12)—his *Ss* expressed the conviction that the series-weights, actually constant, were changed with different anchors.

[12] Stevens, *op. cit.*, this JOURNAL, 71, 1958, 633-646.

[13] The findings of Harvey and Campbell (*op. cit.*, 12-21) comparing changes in weight-estimation in categories-language and in ounces-language, indicate that purely semantic shifts probably do constitute a component of the anchor-effects typically observed with the method of single stimuli.

[14] Wolfgang Köhler, Perceptual organization and learning, this JOURNAL, 71, 1958, 311-315; Wolfgang Köhler and Pauline A. Adams, Perception and attention, this JOURNAL, 71, 1958, 489-503.

[15] Anne Story, Figural after-effects as a function of the perceived characteristics of the inspection figure, this JOURNAL, 72, 1959, 46-56.

[16] A. J. Dinnerstein, Dimensions of stimulation and dimensions of experience. Unpublished Doctoral dissertation, University of Colorado, 1958.

Summary

Anchor-weights, which with the method of single stimuli are known to affect verbal ratings of a weight-series, are here with the method of constant stimuli shown to affect discriminability. Data from four *S*s, who made 2,000 successive comparisons each, show that light vs. heavy anchors, hefted in the non-judging hand, cause contrasting shifts in discriminability within the variable series. This finding supports the view that a type of change in the stimulus-field which can cause shifts in the 'modulus of judgment' would be likely also to cause functionally similar shifts in 'sensory excitability.'

EFFECT OF FLICKER-PERIODICITY UPON PERFORMANCE AND AROUSAL DURING A ROTARY-PURSUIT TASK

By George H. Zimny, Marquette University *

Intermittent photic stimulation, hereafter referred to as flicker, is classified as a stressor agent by Sines *et al.*, and has been used as the stimulus in the investigation of a variety of problems.[1] It is believed by some investigators that aperiodic flicker is more disturbing or stressful than is periodic flicker.[2]

One purpose of the present study was to test this hypothesis by determining the effects of periodic and aperiodic flicker upon performance and arousal during a rotary-pursuit task.[3] A second purpose was to relate performance and arousal. The concept of arousal has been described by a number of authors as referring to the degree of activation and energizing of the organism.[4] It has been found that level of performance is related to level of arousal.[5] Skin-resistance and heart-rate, which Malmo stated were physiological indicants of arousal,[6] were used in the present study. Some question has, however, been raised as to their equivalence in this respect.[7] A third purpose of the present study was to compare heart-rate and skin-resistance as indicants of arousal.

---

* Received for publication March 4, 1964. This research was supported in part by a grant from the Committee on Research of Marquette University.

[1] J. O. Sines, J. A. Stern, L. C. Johnson, and G. A. Ulett, Autonomic correlates of the photically activated *EEG, J. Neuropsychiat.*, 1, 1960, 253-259. *ERDL—Tulane Symposium on Flicker*, 1957.

[2] M. A. B. Brazier, *EEG* studies of flicker in normal man, in L. M. N. Bach (ed.), *Tulane Symposium on Flicker*, 1957, 233 Ernst Simonson, *idem*, 103.

[3] L. M. N. Bach, C. J. Sperry, and J. T. Ray, The effects of flickering light on human subjects, in L.M.N. Bach (ed.), *op. cit.*, 5-51.

[4] Elizabeth Duffy, The psychological significance of the concept of "arousal" or "activation," *Psychol. Rev.*, 64, 1957, 265-275; D. O. Hebb, Drives and the C.N.S. (Conceptual Nervous System), *ibid.*, 62, 1955, 243-254; R. B. Malmo, Anxiety and behavioral arousal, *ibid.*, 64, 1957, 276-287.

[5] A. K. Bartoshuk, Electromyographic gradients as indicants of motivation, *Canad. J. Psychol.*, 9, 1955, 215-230; R. B. Malmo and J. F. Davis, Physiological gradients as indicants of "arousal" in mirror tracing, *ibid.*, 10, 1956, 231-238; R. G. Stennett, The relationship of performance level to level of arousal, *J. exp. Psychol.*, 54, 1957, 54-61; W. W. Surwillo, Psychological factors in muscle-action potentials; EMG gradients, *ibid.*, 52, 1956, 263-272.

[6] Malmo, *op. cit.*, 278-279.

[7] R. C. Davis and A. M. Buchwald, An exploration of somatic response patterns: Stimulus and sex differences, *J. comp. physiol-Psychol.*, 50, 1957, 44-52; J. I. Lacey, Psycho-physiological approaches to the evaluation of psychotherapeutic process and

Method: (1) Subjects. The *Ss* were 24 women randomly selected from a third-semester course in English which was required of all Liberal Arts students.

Apparatus. A rotary-pursuit apparatus with an 11-in. diameter turntable rotating at 45 r.p.m. and a 0.5-in. diameter brass disk located 4 in. from the center of rotation was used.

The flicker apparatus consisted of a 21-in. in diameter plexiglass episcotister painted flat black and mounted on an electric motor rotating at 30 r.p.m. Twelve holes of 1 in. diameter each were cut 9 in. from the center of the episcotister. For periodic flicker, adjacent holes were centered and cut 9.5 in. from center to center; aperiodic flicker, the distances between the adjacent holes were 7, 3, 6, 4, 5, 7, 4, 3, 7, 3, 6 and 5 in., center to center.

A 150-w. GE projector flood lamp was suspended above the path of the holes in the episcotister. A funnel under this path directed the flashes down onto the turntable. The entire episcotister unit was enclosed in a light-tight compartment.

The intensity of ambient lighting provided by a 25-w. bulb was two foot-candles at the turntable. With the episcotister standing still, the maximal intensity of light falling on the turntable through one of the episcotister holes was 20 foot-candles and the minimal intensity, with light from the flood light blocked, was two foot-candles, the ambient value. The subjective impression was that of a very dim light with pronounced flickering.

(2) Measures. Performance-measures consisted of the number of times the stylus-tip hit the disk and the total time the stylus-tip was on the disk during each 1-min. period of task. An electric counter and accumulative timer were used.

Heart-rate, number of beats per minute, was measured by a photoelectric cardiotachometer, developed by Zenz and Mounts, connected to a Beck-Lee Cardi-all.[8]

Skin-resistance was measured by a psychogalvanometer modified to include a 10-turn potentiometer. Continuous adjustment of the potentiometer provided a measure of skin-resistance consisting of units read from the potentiometer. These units multiplied by 20,000 gave the resistance in ohms. A reading was taken every 15 sec. and the mean of the four readings each minute provided the resistance in potentiometer units per minute.

Two types of scores were obtained for heart-rate and skin-resistance. One score consisted of the level of activity. A base-level and a response-level were available for each process. The other score, a 'change'-score, consisted of the difference between the base-level and the response-level. Only the initial base-level was available for computing 'change'-scores at each minute of task.

(3) Procedure. S sat on a stool with the pursuit rotor located at elbow-level. The photoelectric unit was attached to *S*'s left earlobe, and the two electrodes were attached by electrode paste to the tips of the first and middle fingers of the left hand which rested on a padded support. *S* was told that the attachments provided information about physiological responses and that no electric shock would be given.

---

outcome, in E. A. Rubinstein and M. B. Parloff (eds.), *Research in Psychotherapy*, 1953, 197; G. H. Zimny, Effects of music upon *GSR* and heart rate, this JOURNAL, 76, 1963, 311-314.

[8] Carl Zenz and Frank Mounts, Determination of heart-rate during work, *Arch. indus. Health*, 7, 1958, 280-286.

After the pursuit-instructions had been given and the pursuit demonstrated, *S* was told that a delay was necessary to calibrate the recording instruments operated by an assistant in another room. *E* sat down behind *S* at a small desk containing the timer, counter, and switches for operating the pursuit-rotor and episcotister. During the several minutes of delay, the assistant recorded *S*'s basic heart-rate and skin-resistance. When 'calibration' was completed, *E* started the pursuit-rotor and episcotister, and *S* began his pursuit. The task lasted 5 min. without a rest-period.

The *Ss* were randomly divided into two groups of 12 each. One group received aperiodic flicker and the other, periodic flicker during the entire task. The *Ss* in the two groups were run individually in alternating order.

Results: (1) Performance. Because of the high correlation between total time-on-target and total number of hits for each group (periodic $r = 0.76$,

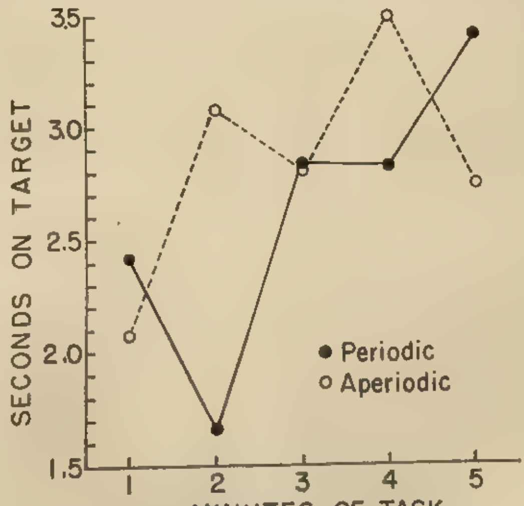

FIG. 1. EFFECT OF FLICKER PERIODICITY UPON PERFORMANCE

$P < 0.01$; aperiodic $r = 0.98$, $P < 0.01$), analysis of performance was made only with the time-measures. Time-on-target during each 1-min. period of the task is shown in Fig. 1. As the analysis of trends for time-on-target showed,[9] there was no significant over-all mean difference between the two experimental conditions. The significant 'minutes-effect' ($F = 2.51$, $P < 0.05$) indicates, however, that the average slope of the two trend lines was significantly different from zero and, therefore, that the performance of the combined groups showed some improvement dur-

[9] Don Lewis, *Quantitative Methods in Psychology*, 1960, 379-398.

ing the task. The significant interaction between minutes and conditions ($F = 2.55$, $P < 0.05$) indicates that the two trend lines were not parallel, but, as is evident from Fig. 1, performance under one condition was not consistently superior to that under the other condition. The largest difference, that at the 2-min. period, was not significant ($t = 1.79$, $P > 0.05$), thus emphasizing the importance of the trend-difference between the two conditions. Because of the crisscrossing of the trend lines, for which no explanation is available, little importance is attached to this nonparallelarity of trend.

(2) Level of skin-resistance. The mean and standard deviation of the basic skin-resistance, measured prior to the onset of the task, were 3.07

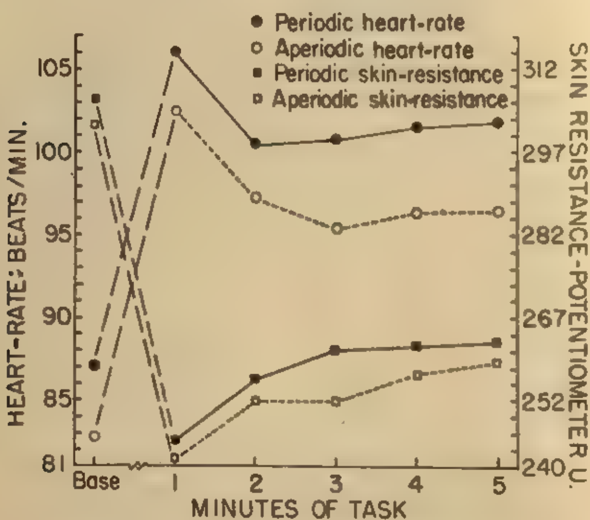

FIG. 2. HEART-RATE AND SKIN-RESISTANCE DURING THE TASK

and 1.55 for the periodic condition and 3.02 and 0.90 for the aperiodic condition. Neither the means ($t = 0.13$, $P > 0.80$) nor variances ($F = 2.96$, $P > 0.05$) were significantly different.

The mean level of skin-resistance for each minute of task was computed for each S from the four readings obtained during each minute. Fig. 2 (right ordinate) shows this mean for each minute of task. The over-all mean difference between the two groups was not significant. The significant 'minutes-effect' ($F = 10.42$, $P < 0.01$) indicates that the average slope of the two trend lines was significantly different from zero. The

Tukey gap-test of differences between successive minutes reveals that the skin-resistance at the first minute was significantly lower than at the remaining four. For the combined groups, a significant difference in means (correlated groups $t = 4.68$, $P < 0.01$) but not variances ($F = 1.76$, $P > 0.10$) existed between the base-level and the highest level at any one period—the fifth minute. The absence of a significant interaction between 'minutes' and 'conditions' indicates parallelarity of the two trends. Thus, a significant decrease from the base-level occurred during the first minute of task followed by a significant rise in skin-resistance to a level that was consistent during the remaining four minutes of task but significantly below the base-level.

(3) Level of heart-rate. The mean and standard deviation of the base heart-rate, measured prior to the onset of the task, were 87.08 and 11.44 for the periodic condition and 82.83 and 15.37 for the aperiodic condition. Neither the means ($t = 0.75$, $P > 0.40$) nor variances ($F = 1.90$, $P > 0.10$) were significantly different.

The mean level of heart-rate (left ordinate) for each minute of task is also shown in Fig. 2. The conclusions based on the results of the trend analysis and the Tukey gap-test for heart rate are identical to those for skin-resistance. For the combined groups, a significant difference in means (correlated groups $t = 8.57$, $P < 0.01$) but not variances ($F = 1.13$, $P > 0.10$) existed between base-level and the lowest level at any one period—the third minute. Thus, the pattern of response was the same, except inverted, for level of heart-rate as for level of skin-resistance.

(4) 'Change'-scores and the law of initial values. Prior to any analysis using 'change'-scores, correlation coefficients between base-level and amount of change from base-level for the combined groups ($N = 24$) were computed to investigate the possible operation in the study of the so-called law of initial values. The product-moment coefficients at the first through the fifth minute of task were 0.82, 0.80, 0.77, 0.75, and 0.62, respectively, for skin-resistance and −0.27, −0.25, −0.13, −0.02, and −0.04, respectively, for heart-rate. For skin-resistance, all coefficients had $P < 0.01$. For heart-rate, all coefficients had $P > 0.05$. These results indicate that the law of initial values was operating in the case of skin-resistance but not heart-rate.

In an effort to remove the influence of base-level upon the magnitude of change from base-level in the case of skin-resistance, a relative measure consisting of the amount of change in resistance from the base-level divided by the base-level was computed for each S at each 1-min. period. The effectiveness of this procedure is indicated by the product-moment

coefficients between base-level and the relative measure of change from base-level. The coefficients for each successive minute of task were 0.37, 0.39, 0.36, 0.34, and 0.36. All coefficients had $P > 0.05$.

(5) Changes in skin-resistance. The trend analysis was repeated using the relative measure of change in skin-resistance. The conclusions are the same as those reached for (2), level of skin-resistance.

(6) Changes in heart-rate. The trend analysis was carried out for heart-rate, with the amount of increase in heart-rate over the base-rate as the change-score. The conclusions reached are the same as those for (3), level of heart-rate.

(7) Relations among measures of response. Since no significant differences existed between the two experimental groups, they were combined in computing the product-moment correlations given below. No significan correlation was found between base-level of heart-rate and base-level of skin-resistance, ($r = -0.07$, $P > 0.05$). The correlations between the maximal response of the two physiological processes during the task were: 0.04, $P > 0.05$, between levels of heart-rate and skin-resistance; 0.04, $P > 0.05$, between changes of heart-rate and skin-resistance; and 0.20, $P > 0.05$, between relative changes of heart-rate and skin-resistance. The correlations between the maximal physiological response during the task and maximal time-on-target were: 0.36, $P > 0.05$, between level of heart-rate and time; 0.15, $P > 0.05$, between level of skin-resistance and time; 0.53, $P < 0.01$, between change in heart-rate and time; 0.13, $P > 0.05$, between change in skin-resistance and time; and 0.06, $P > 0.05$, between relative change in skin-resistance and time. Thus, there were no significant correlations between the two physiological responses and only one significant correlation between a physiological response and performance; namely, the positive correlation between maximal heart-rate change during task and maximal time-on-target.

Discussion: Flicker periodicity. The results clearly indicate that the expected differential effects of periodic and aperiodic flicker upon performance and arousal were not obtained, although both measures of arousal did indicate that the task was arousing. It should be noted, however, that the expectation of greater interference due to aperiodic than to periodic flicker was stated in relation to *EEG* alpha rhythm by Brazier and to the perception of small moving objects by Simonson.[10]

Arousal and performance. As Duffy points out and Woodworth and

[10] Brazier, *op. cit.*, 207; Simonson, *op. cit.*, 103.

Schlosberg imply, there is no agreement on the basic principles underlying the relationship between level of arousal during a task and level of performance.[11] The lack of agreement probably stems in large part from the fact that the level of arousal can be influenced by quite a number of factors including: (a) the activity connected with the task; (b) *S*'s motivation to carry out the task; and (c) his anxiety and apprehension at doing the task. In the present study, a marked increase in the level of arousal accompanied the onset of the task, but the exact source of the increase cannot be identified because all the factors mentioned above could have been operating. The increase in performance-level during the last four minutes of the task was not accompanied by any change in arousal-level during that period but, again, no exact interpretation of the observed relationship is possible because of the varying ways in which the previously cited factors, plus the additional ones of practice and fatigue could have interacted. The shape of the trends of arousal and performance found here appear to be similar to those found by Kling, Williams, and Schlosberg for the initial five minutes of the rotary-pursuit task used in their study.[12] Thus, the general relationship between arousal and rotary-pursuit found in these two studies is initially high arousal and low performance followed by improving performance and constant arousal.

Another approach to the relationship between arousal and performance is by means of correlations between measures of physiological processes and performance-scores. The few studies that have employed this approach have produced conflicting results.[13] In the present study, the only significant correlation was that of 0.53 between maximal change in heart-rate during the task and maximal time-on-target. Because an inverted U relationship between arousal and performance has been postulated, the scatter-plot for each correlation was examined, but no indication of curve-linearity was noted.[14] Finding that only one measure of one physiological process was related to performance indicates that differences among physiological responses is a complicating factor.

Physiological measures of arousal. Both heart-rate and skin-resistance yielded the same conclusions concerning arousal; namely, that arousal was

[11] Elizabeth Duffy, *Activation and Behavior*, 1962, 156; R. S. Woodworth and Harold Schlosberg, *Experimental Psychology* (rev. ed.), 1954, 146-148 and 175-177.

[12] J. W. Kling, J. P. Williams, and Harold Schlosberg, Patterns of skin conductance during rotary pursuit, *Percept. mot. Skills*, 9, 1959, 303-312.

[13] Malmo and Davis, *op. cit.*, 235; Surwillo, *op. cit.*, 268.

[14] Duffy, *op. cit.*, 268; G. L. Freeman, The relationship between performance level and bodily activity level, *J. exp. Psychol.*, 26, 1940, 602-608; Hebb, *op. cit.* 250-251; Malmo, *op. cit.*, 277-278; Stennett, *op. cit.*, 54.

greatest during the initial minute of task, then dropped to a lower but still above base-level, and remained there during the last four minutes. Such perfect agreement in conclusions from different indicants of arousal is sometimes found.[15] Lack of agreement is, however, considerably more common.[16] Although Lacey and Schnore have indicated that level-scores and change-scores of physiological activity are different in several respects, still, in the present study, the same conclusions concerning arousal were reached from both types of scores for each physiological process.[17]

The consistency of these conclusions might appear to indicate that heart-rate and skin-resistance were occurring in parallel fashion within each *S*. None of the correlations, however, between heart-rate and skin-resistance for base-level, for level during task, and for change from base during task was significantly different from zero. Thus, there is pronounced intra-individual variability in the two processes, a finding which is in line with the concept of fractionation or idiosyncratic patterning of physiological processes.[18]

Further evidence of the difference in response of the two physiological processes is provided in the present study by the finding that the law of initial values was operating in the case of skin-resistance but not heart-rate. No explanation can be given as to why the response in one process is related to the base-level whereas the response in the other is not.

Summary. Each of 24 *S*s, college women, carried out a rotary-pursuit task for 5 min. under either a periodic or an aperiodic flickering light while measures of performance and arousal were obtained from them. The flicker conditions had no differential effects upon either performance or arousal. Performance improved during the task. All measures of heart-rate and skin-resistance indicated that arousal was greatest during the initial minute of the task, then dropped to a lower level, but still above the base, and remained there for the duration of the task, *i.e.*, the last 4 min. Differences between heart-rate and skin-resistance as indicants of arousal were also considered.

---

[15] Stennett, *op. cit.*, 60.

[16] Davis and Buchwald, *op. cit.*, 46; S. M. Feldman, Differential effects of shock in human maze learning, *J. exp. Psychol.*, 62, 1961, 171-178; Lacey, *op. cit.*, 197; M. M. Schnore, Individual patterns of physiological activity as a function of task differences and degree of arousal, *J. exp. Psychol.*, 58, 1959, 117-128; Zimny, *op. cit.*, 313.

[17] Lacey, *op. cit.*, 179; Schnore, *op. cit.*, 121.

[18] Lacey, *op. cit.*, 190-191; Schnore, *op. cit.*, 118.

PROBABILITY-DISCRIMINATION IN THE FISH

By Erika R. Behrend, Barbara A. Bauman and M. E. Bitterman,
Bryn Mawr College

In a recent experiment with the pigeon, latency of response in a single-key situation was found to be affected by probability of reinforcement, but not by frequency of reinforcement.[1] With a 70% probability of reinforcement for response when the key was one color, a 30% probability of reinforcement for response when the key was a different color, and an equal number of reinforcements earned for response to the two colors, latency of response was lower for the 70% color than for the 30% color; with 70% of reinforcements earned for response to one color, 30% of reinforcements earned for response to a different color, and a 50% probability of reinforcement for response to each color, latency of response to the two colors was the same. The one-key experiment was prompted by an interest in the basis of the so-called probability-matching which the pigeon shows in certain choice- (two-key) situations under conditions in which probability and frequency of reinforcement are confounded.[2] The one-key results do not, of course, unequivocally establish probability as the basis of matching in the two-key situation, since the determinants of latency and of choice may be quite different, but there is no other basis for a decision until the two variables can be unconfounded in a choice-situation. We report here a one-key experiment with the fish *Tilapia macrocephala* (the African mouthbreeder), which also has been found to match in choice-experiments.[3]

Method

Subjects. The *Ss* were nine sexually mature mouthbreeders, ranging in length from 3–4 in., which had been bred in the laboratory. Five of the nine had previously been used in a one-key experiment on partial reinforcement, and the remaining four in a two-key experiment on habit-reversal.

* Received for publication June 24, 1963. This research was supported by Grants MH-02857 and MH-7683 from the Public Health Service.

[1] Virgil Graf, D. H. Bullock, and M. E. Bitterman, Further experiments on probability-matching in the pigeon, *J. exp. Anal. Behav.*, 7, 1964, 151-157.

[2] Bullock and Bitterman, Probability-matching in the pigeon, this Journal, 74, 1961, 634-639.

[3] M. E. Bitterman, Jerome Wodinsky, and D. K. Candland, Some comparative psychology, this Journal, 71, 1958, 94-110; E. R. Behrend and M. E. Bitterman, Probability-matching in the fish, this Journal, 74, 1961, 542-551.

Apparatus. The apparatus employed was a modification of that described by Longo and Bitterman.[4] It consisted of (1) a darkened enclosure of black Plexiglas to which each animal was brought in its individual 2-gal. living tank; (2) a target—a thin disk of diffusing Plexiglas—mounted on the lid of the enclosure, which was lowered into the water at one end of the tank as the lid was brought down; (3) a light-box, containing Christmas-tree lamps, for projecting lights of different color on the target; (4) a pellet-dispenser; and (5) equipment for programming discrete trials of several different kinds, for detecting responses, and for measuring response-latencies and rates. Response-detection was accomplished by an amplifier into which was fed the output of a crystal phonograph-cartridge, whose needle-holder supported a thin rod to which the target was fixed.

Procedure. Each trial began, after a 6-sec. intertrial interval in darkness, with the illumination of the target by a red or a green lamp, at which instant a Standard Electric 0.01-sec. timer was activated. Response stopped the timer, turned off the target-light, and (on unreinforced trials) began the intertrial interval, or (on reinforced trials) initiated the reinforcement cycle. Reinforcement was a pellet of dry food dropped into the water at the rear of the tank (the end opposite the target); as the pellet was dropped, the inner rear wall of the enclosure (a piece of milk-Plexiglas) was illuminated for 6 sec. by a white lamp behind it, permitting the fish to find the food. On some trials (see below), 'response' was defined not as one, but several (as many as 10) strikes at the target. On such trials, two timers were activated, one to measure latency of the first response, and the second to measure the time required to run the ratio. If on any trial, an animal failed to meet the response-requirement in 40 sec., the trial was terminated, and then repeated at the close of the session.

There were three main experimental conditions: *confounded*, *equal-probability*, and *equal-frequency*. In the confounded condition, 20 trials per day were given, 10 with red and 10 with green light, in Gellermann orders.[5] Of the 10 trials with one color (say, red), 7 were reinforced and 3 were not; of the 10 trials with green, 3 were reinforced and 7 were not. The pattern of reinforcement was random, with the restriction that there be no more than one red nonreinforcement and no more than one green reinforcement in each block of five trials. In the equal-probability condition, the ratio of reinforcement was 70:30, but the probability of reinforcement for response to each color was 50%. For example in each daily series of 20 trials, there might be 7 reinforced responses to red, 7 unreinforced responses to red, 3 reinforced responses to green, and 3 unreinforced responses to green. In the equal-frequency condition, the number of reinforced responses to the two colors was equated, while the probability of reinforcement was about 70% for response to one color and about 30% for response to the other. For example, in each daily series of 24 trials, there might be 5 reinforced responses to red, 2 unreinforced responses to red, 5 reinforced responses to green, and 12 unreinforced responses to green. In the equal-probability and equal-frequency conditions, as in the confounded condition, the various kinds of trial were presented in quasi-random, balanced orders.

---

[4] Nicholas Longo and M. E. Bitterman, Improved apparatus for the study of learning in fish, this JOURNAL, 72, 1959, 616-620.

[5] L. W. Gellermann, Chance orders of alternating stimuli in visual discrimination experiments, *J. genet. Psychol.*, 42, 1933, 206-208.

As in previous work with the pigeon under analogous conditions, the confounded problem did not produce any substantial difference in latency of response to the two colors for most of the fish when response was defined as a single strike at the target. Accordingly, an *FR*-schedule was introduced, with the number of strikes required to terminate each trial being increased gradually. As it had in pigeons, this procedure produced differential response to the two colors in all of the mouthbreeders, although the ratios required (usually 2-5, and never more than 10) were much lower for the mouthbreeders than for the pigeons. Another difference between the two species is that, in the mouthbreeders, the *FR* produced differential response as measured both by latency and by ratio-time. While the ratio was necessary to produce differential response in the pigeons, the discrimination was evident only the latency of the initial response on each trial—once the ratio got under way, it was run off at a fairly constant rate. The difference probably is to be understood in terms of the motor properties of the two species.

Most of the fish studied were started on the confounded problem and then transferred to one or another of the remaining two conditions, or to both in different orders. A few fish began with the equal-frequency condition.

Results

The results for fish No. 11, a representative animal trained under each of the three conditions, are shown in Fig. 1. The curves are plotted in terms of mean log latency and mean log ratio-time for each stimulus. (The logarithmic transformation was used to normalize the data, since latency-scores tend to be skewed positively. A constant of +2 was added to each log-score to avoid the necessity of working with negative numbers.) There were five stages in the training of this animal which are set off from each other in the figure by vertical lines.

The first portion of Fig. 1 (Days 1-9) shows the performance of this animal under the confounded condition with *FR*4 and green as the 70% color. The animal's training under this condition had begun without the ratio and it was already showing lower latency of response to green when the ratio was introduced. With the ratio, the animal discriminated in terms of both measures.

In the second stage of this animal's training, the equal-probability condition was used, with green again as the majority (more frequently reinforced) color. As may be seen in the second portion of Fig. 1, the differences in mean latency and in mean ratio-time became very small, although they continued to favor the more frequently reinforced stimulus. In the third stage of the experiment, an attempt was made to determine whether these differences were due to the greater frequency with which green was reinforced in the second stage, or whether they could be regarded simply as a residual effect of the preceding confounded training; the majority and minority colors were reversed in frequency while the probability remained equal (seven reinforcements and seven nonreinforcements for response to

green). The third portion of Fig. 1 shows no substantial difference in response to the two colors in the third stage; the only sign of a difference, which appeared on Days 32 and 33, favored green (the majority stimulus of the first stage) which was now the *less* frequently reinforced. Under these conditions, then, the fish does not seem to discriminate frequency of reinforcement.

In the fourth stage, the animal was shifted to an equal-frequency condition with red—the more frequently reinforced color of the third stage—

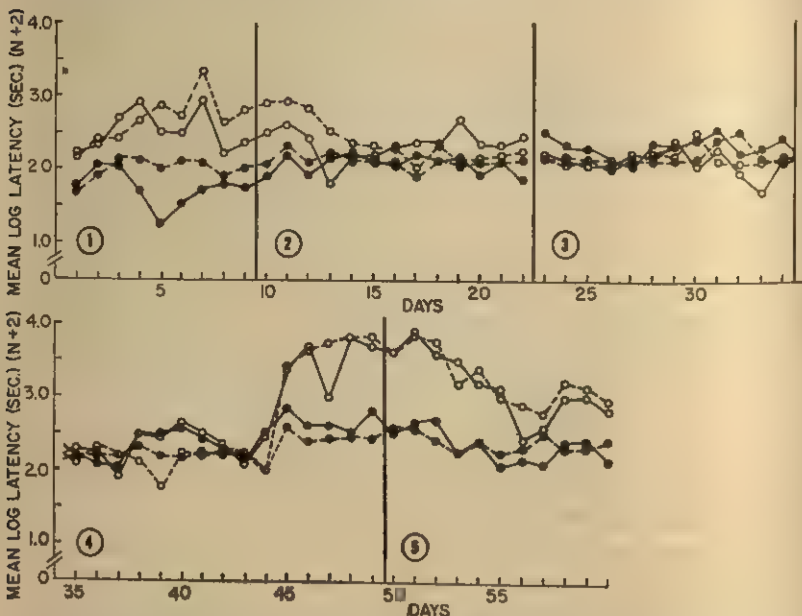

FIG. 1. THE DATA FOR A REPRESENTATIVE FISH (NO. 11)
The solid lines are plots of latency and the broken lines are plots of ratio-time. The open circles refer to the minority stimulus while the closed circles refer to the majority stimulus. The conditions of training in each stage are as follows: 1, confounded; 2 and 3, equal probability; 4, equal frequency; 5, equal probability. These conditions are described in the text.

now the majority color in probability. As is shown in the fourth portion of Fig. 1, there was no sign of discrimination for about 10 days, but thereafter the evidence of discrimination became quite clear. In the fifth stage of its training, the fish was shifted once more to the equal-probability condition with red as the more frequently reinforced color. As the final portion of Fig. 1 illustrates, the discrimination began to decline, although it was necessary to terminate the experiment before it could be determined

whether differential response to the two stimuli would disappear altogether.

The results for the other animals were very much like those of No. 11 with one exception: while they differentiated sharply in terms of both measures under the equal-frequency condition as well as under the confounded condition, and did not differentiate in terms of latency under the equal-probability condition, there was some evidence of differentiation in terms of rate under the equal-probability condition. The results for the equal-frequency and equal-probability conditions must now be considered in detail.

To obtain a quantitative index of the consistency of the discrimination made by each animal under each condition, we computed separately, for latency and for ratio-time the daily mean rank of the responses to the minority stimulus for comparison with the mean rank to be expected by chance and the mean rank which would result from perfect discrimination. In each case, the rank expected by chance would be the mean of a random sample of ranks from 1 to N (the number of trials given), or $\frac{1}{2}(N+1)$. With perfect discrimination, the response-times for the minority stimulus would be greater in every case than the response-times for the majority stimulus, and the mean minority rank, therefore, would be the mean of the M highest ranks (where M is the number of minority trials).

Seven animals were tested under the equal-frequency condition. The mean minority rank expected by chance would be 12.5, since there were 24 trials per day, while perfect discrimination would give a mean minority rank of 16 (the average of ranks 8-24). The mean of the minority ranks for the last three days of training was 14.3 for latency and 14.9 for ratio-time. The corresponding ranges were 11.9-15.6 and 11.4-16.0. The lower limits of these ranges reflect the performance of the only one of the seven animals that did not reliably discriminate probability.

Testing under the equal-probability condition yielded mixed results. In Stage 2, the two other animals which were tested under the same conditions as No. 11 also continued to respond to the majority stimulus with lower latency and ratio-time. In Stage 3, when the majority stimulus for frequency was reversed, they continued, like No. 11, to respond with lower latency to the previous majority stimulus, but, unlike No. 11, they began to respond to the new majority stimulus with lower ratio-time. The mean minority rank expected by chance would be 10.5 (since there were 20 trials per day), while perfect discrimination would be indicated by a score of 17.5 (the average of ranks 15-20). The mean minority ranks for ratio-time on the last three days of Stage 3 was 9.3 for No. 11; for the two other animals the means were 12.9 and 13.4. These fish were tested again

under the equal-probability condition after experience with the equal-frequency condition, as were three other fish not previously tested with equal-probability. All the animals, which differentiated markedly in equal-frequency, showed decline in differentiation in equal-probability. Unfortunately, there was not sufficient time to carry the animals to asymptote. The question of whether the animals are affected by frequency in this situation cannot, then, be answered with confidence. About the effect of probability independently of frequency, there is, however, no doubt at all.

Discussion

The present results for the fish are identical, insofar as *latency* is concerned, to those obtained in an analogous experiment with the pigeon: both species discriminate probability when frequency is equated but fail to discriminate frequency when probability is equated. We have no intention, of course, to deny the possibility of frequency-discrimination in either species on the basis of these results. It would not be surprising, for example, if a more drastic variation in frequency produced differential response. We only maintain that, over the range of frequencies studied in the one-key experiments, and in the choice-experiments which the one-key experiments were designed to help us understand, probability is a far more powerful variable than frequency.

Insofar as *rate* (ratio-time) is concerned, the present results for the fish are different from those for the pigeon. The pigeon's ratio-times were sensitive neither to probability, nor to frequency, nor to the two variables together. In the fish, by contrast, ratio-time was markedly influenced by probability, and to some extent also by frequency. We probably are dealing here with a motor difference; the facility with which the pigeon generates high rates is so great that much higher ratios than those used in the pigeon experiment might be necessary to detect a discrimination of frequency.

Whatever the reason for the difference in the results for fish and pigeon, the results for both species are in accord in suggesting that probability is a more powerful variable than frequency. These experiments provide some support, then, for the assumption that the matching which both species show in the two-key situation is based on a discrimination of probability rather than of frequency. We cannot, of course, be sure that the determinants of performance in the two types of situation are the same, but a parallel experiment with the rat might be instructive. If the rat, which does not match in the two-key situation,[6] behaved differently also in the one-key

[6] Bitterman, Wodinsky, and Candland, *op. cit.*, 105-107.

situation, there would be some further ground for the belief that the determinants of behavior in the two situations are the same. If the one-key results for the rat were the same as for the fish and the pigeon, we would have to conclude that the determinants of behavior in the two situations are different.

SUMMARY

In a discrete-trials, one-key experiment with the African mouthbreeder, probability and frequency of reinforcement were varied independently. Differences in probability produced marked differences in latency and in rate of response. Differences in frequency had no effect on latency of response and only a small effect on rate. The results are considered in relation to those obtained in an analogous experiment with the pigeon, and to those obtained in choice-experiments with both species.

THE INFLUENCE OF LIGHTNESS UPON THE JUDGMENT OF SIZE

By WILLIAM BEVAN and EDWARD D. TURNER, Kansas State University

Traditional psychophysics is concerned with the delineation of dimensional relationships between stimuli and responses with variation on all stimulus-dimensions save one removed, but frame-of-reference psychophysics has demonstrated that judgments even of the simplest stimulus-properties are the product both of focal and of contextual variables. As early as 1938, Richardson pointed out the need for a multidimensional psychophysics and proposed a method for the scaling of stimulus-similarity which resembles factor analysis.[1] In 1950, Attneave reported an elaborate study of similarity among forms differing multidimensionally.[2] He concluded that judgments are composites of the sort to be predicted by a multiple regression-equation, and that situational variables like spatial arrangement of forms are important factors in judgment.

The past decade has seen the development of the multidimensional scaling approach, which involves a Euclidean analysis of psychological distance derived from judgments of similarity. Multidimensional scales have been produced for many properties of stimuli. For example, Ekman and Engen reported a scaling of two ordorants and their mixtures,[3] and Abelson and Sermat, a scaling of facial expressions.[4] Recently, Messick has reviewed theoretical developments in multidimensional scaling.[5]

The present exercise explores an alternative approach to multidimensional psychophysics. A unidimensional scaling method is employed, and dimensional interaction is assessed through manipulations within the experimental design. The method is proposed for multidimensional situa-

* Received for publication August 15, 1963. This experiment was performed under contract Nonr-3634(01) between Kansas State University and the Office of Naval Research.

[1] M. W. Richardson, Multidimensional psychophysics, *Psychol. Bull.*, 35, 1938, 659-660.

[2] Fred Attneave, Dimensions of similarity, this JOURNAL, 63, 1950, 516-556.

[3] Gösta Ekman and Trygg Engen, Multidimensional ratio scaling and multidimensional similarity in olfactory perception, *Reports from the Psychological Laboratory, University of Stockholm*, No. 126, May, 1962, 1-11.

[4] R. P. Abelson and Vello Sermat, Multidimensional scaling of facial expressions, *J. exp. Psychol.*, 63, 1962, 546-554.

[5] S. J. Messick, Some recent theoretical developments in multidimensional scaling, *Educ. Psychol. Meas.*, 16, 1956, 82-109.

tions in which a specifiable number of identifiable stimulus-variables are to be studied. Unlike the traditional approach, which excludes all independent variables except one, it provides for judgments on a criterion-variable, while this latter stands, in principle, in all degrees of relationship with one or more incidental variables.

Method

This experiment deals with the simple illustrative case of one incidental variable. The response called for is judgment of size. The criterion-variable is physical size, and the incidental variable is lightness of patch. The general procedure is as follows: Three stimulus-series are prepared in which the variations on the criterion- and incidental variables are uncorrelated (the relationship maintained in traditional psychophysical experiments) and perfectly correlated, both positively and negatively. If empirical *S–R* functions are then obtained for the last two conditions, they define the limits of a two-dimensional space within which judgments may be predicted for any intermediate degree of relationship between the criterion- and incidental variable. By determining the mathematical properties of the curve of best fit for the two extremes of relationship, it is possible to construct a predictive function for any intermediate degree of relationship. For example, if the *S–R* functions are linear, a plot of the slopes of the best fitting lines for *r*s of 1.00, 0.00, and −1.00, will yield an estimate of the slope of the *S–R* function for any intermediate degree of relationship.

Subjects. A total of 59 undergraduate students served as *S*s. Three groups of 13 *S*s each were used under the 1.00, 0.00 and −1.00 conditions, respectively, to derive predictions for the remaining two groups of 10 *S*s tested with correlations of 0.60 and −0.60, respectively. There were approximately equal numbers of men and women in each group.

Materials. All stimuli were squares cut from Munsell grays and mounted on a white background. There were five series of five cards each. In size, the squares were 44, 47, 50, 53, and 56 mm. on a side. In the uncorrelated series, all were the same intermediate gray (Munsell No. 5). In the correlated groups, value varied from No. 3 through No. 7, with no two patches having the same value. In the 1.00 series, as size increased, lightness also increased. In the −1.00 series, as size increased, lightness decreased. In the 0.60 series, as size increased, lightness varied as follows: Munsell No. 3, 6, 5, 4, 7. In the −0.60 series, as size increased, lightness varied as follows: 7, 4, 5, 6, 3.

Procedure. The psychophysical method employed was a variation of the Newhall Method used in earlier studies of hysteresis.[8] The series-members were presented in sequence from smallest to largest as a series of overlapping pairs, with the larger member of each prior pair becoming the smaller member of each subsequent pair. Each pair was presented in a Gerbrands Harvard two-field tachistoscope, the smaller exposed first and immediately followed by the larger. The exposure-time for each member of the pair was 0.5 sec.

S was requested to represent the size of the second square relative to that of the

[8] William Bevan, Harry Barker, and J. F. Pritchard, The Newhall scaling method, psychophysical bowing, and adaptation level, *J. gen. Psychol.*, 69, 1963, 95-111.

first by appropriately locating a marker on a grid placed before him. Along the abscissa of the grid were placed 5 small steel washers, equally spaced to represent the 5 members of the stimulus-series. The position of the left-most washer, representing the smallest series-member, was fixed at zero on the ordinate. All other markers, including that representing the largest size, were free to be moved. The ordinate was arbitrarily divided into 100 equal-sized units, but it was made clear to *S* that he need use only as much of the scale as he deemed necessary. He also was free to add units if he considered it appropriate to do so. Thus, when *S* had judged the relative size of the final pair of the series, he had constructed a graphic function of judged size. The series was repeated twice and the median of the three settings for each stimulus was taken as representative of *S*'s performance.

RESULTS

The relation of brightness to judged size. The mean data of the first three groups are shown in Fig. 1. It is clear (between conditions $F = 5.42$, $df = 2/36$, $p < 0.01$) that stimulus-lightness influences judged size, but any overall generalization about the influence of the former upon the latter must include a consideration of the systematic relationship between lightness and physical size across the stimulus-series. The common textbook-statement that lightness enhances apparent size lacks adequate precision. Lightness enhances apparent size when lightness and size are positively related as properties of the stimulus-series. The converse is true when the correlation is negative.

Of particular interest is the fact that the curves for the correlation-conditions display hysteresis, positive correlation yielding bowing upward and negative correlation bowing downward. Stevens has associated this, in the method of bisection, with ascending and descending orders of presentation, respectively,[7] while Bevan, Barker, and Pritchard, using the Newhall Method, have found it to be related to the direction of magnitude-change in the individual pairs.[8] The present data suggest that order displays its influence through incidental as well as criterion-variables, and it is especially noteworthy that, in the case of negative correlation, the order of the former seems to overrule that characterizing the latter. The explanation of upward and downward bowing in terms of shifting *AL*s, proposed by Bevan, Barker, and Pritchard, would also appear to be appropriate in this case.

Judgment with intermediate degrees of relationship. The second phase of the experiment involved testing the predictive power of this simple model by obtaining data for intermediate degrees of relationship ($\pm$ 0.60) between criterion- and incidental variables. Since judgments on members

[7] S. S. Stevens, On the psychophysical law, *Psychol. Rev.*, 64, 1957, 153-181.
[8] Bevan, Barker, and Pritchard, *op. cit.*, 95-111.

of successive series obtained with the Newhall method, are cumulative and therefore correlated, since the intensitive steps are not equal on successive trials, and since the correlation between the two stimulus-variables changes

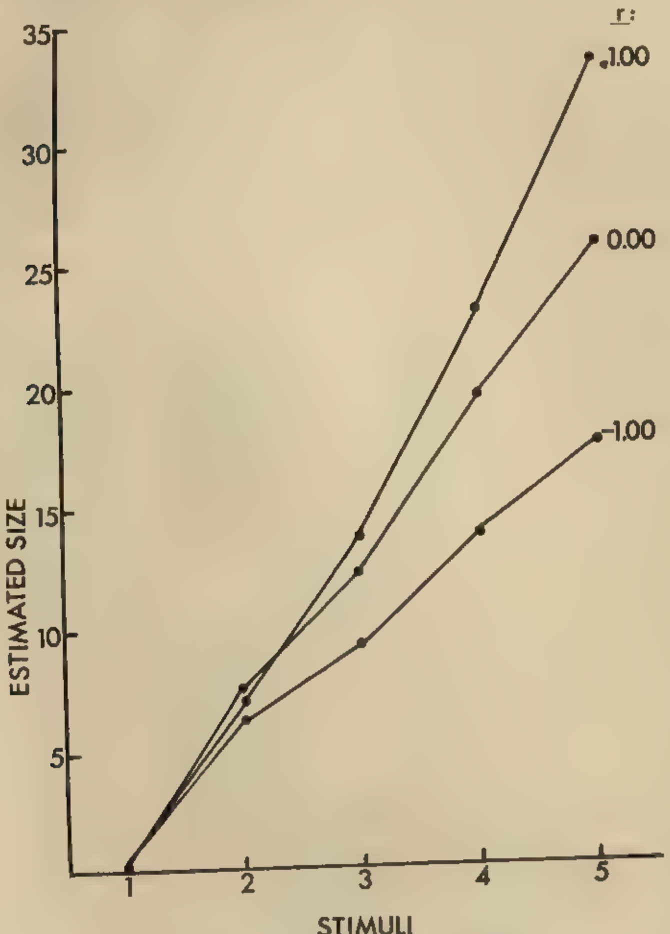

FIG. 1. JUDGMENTS OF SIZE BY THE NEWHALL METHOD
When an incidental variable (lightness) is positively correlated, negatively correlated, or uncorrelated with size in the stimulus-series.

on successive trials, being ± 0.60 only on the last trial, prediction is quickly checked for the final series-member. Assuming that the regression of the judgments of size upon color is linear, prediction for conditions of positive correlation are obtained by determining the difference in response-

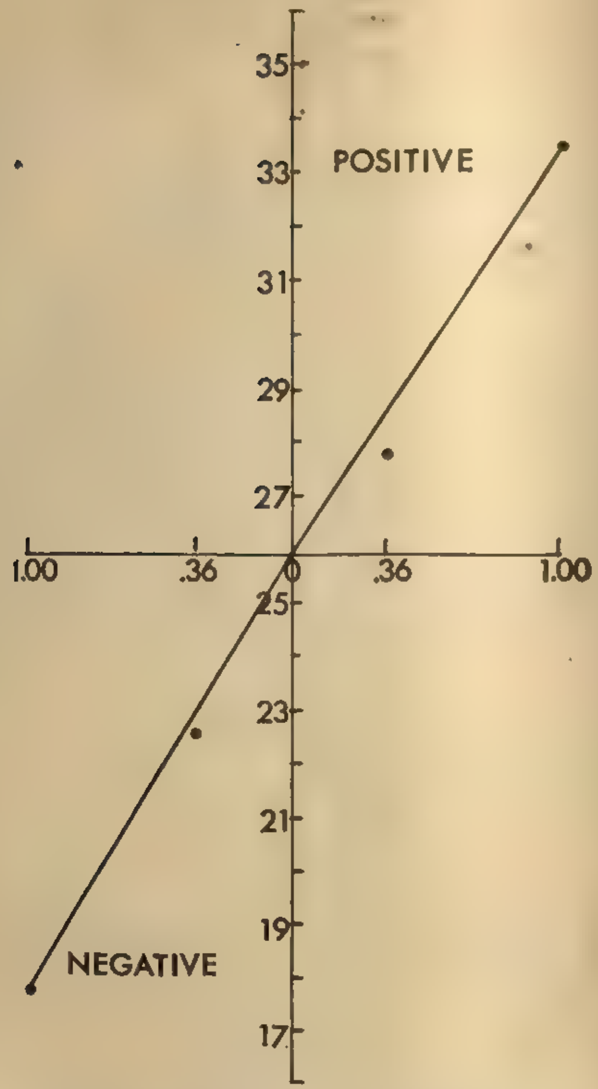

FIG. 2. PREDICTION-MATRIX FOR STIMULUS-FIVE WITH ALL DEGREES OF RELATIONSHIP BETWEEN CRITERION AND INCIDENTAL VARIABLES

The values plotted were those obtained for positive and negative correlations of 0.60.

scale settings for each series-member under the conditions zero and perfect correlation, multiplying these values by r^2 (in this case, 0.36), and adding the product to the scale-value for the zero condition. For the conditions of negative correlation, the procedure is the same, except that the final operation is subtraction rather than addition. The predicted value for the 0.60 condition is 28.6, and the obtained value 27.7; the error is 3%. The predicted value for the −0.60 condition is 22.9, and the obtained value is 22.5; the error is 2%. Agreement thus is rather good.

Predictions also may be obtained graphically, of course, by plotting response-scale values against r^2. A prediction-matrix for Stimulus Five, shown on Fig. 2, indicates the rather good agreement between estimated and obtained response-values just noted.

Summary

The traditional approach to the study of psychophysical relationships has been to isolate the stimulus-variable under consideration from all other stimulus-properties. The limitations of this procedure for an understanding of judgment under complex stimulus-conditions is readily apparent. An alternative approach is suggested in which psychophysical judgments are made when the variable to be judged stands in a prescribed relationship to one or more additional stimulus-variables. The special case of one incidental variable is experimentally examined. By obtaining judgments when the variable to be judged (criterion-variable) and the incidental variable are perfectly correlated, positively and negatively, as well as for the traditional condition under which the incidental variable is held constant, it is possible to construct a space that will provide for the prediction of judgments under all degrees of relationship between the two variables.

The variable chosen for judgment was the size of square gray patches mounted on a white background. The incidental variable was lightness of patch, the series members varying from dark to light gray. The psychophysical method was that of Newhall. When the correlation between lightness and size was 1.00, size-estimates were enhanced; when the correlation was −1.00, the effect was one of diminution. Under the condition of positive correlation, the Newhall function was bowed upward, while under the condition of negative correlation, the opposite was the case. Judgments obtained for stimuli terminating series with an intermediate degree of positive or negative relationship showed good agreement with predicted values.

THE EFFECT OF DEMAND CHARACTERISTICS ON THE FIGURAL AFTER-EFFECT WITH REAL AND IMAGED INDUCING FIGURES

By GEORGE SINGER and P. W. SHEEHAN, University of Sydney, Australia

Following protracted visual stimulation by a contour, a subsequently presented contour slightly removed from the first is judged as displaced or distorted. Such spatial displacements are referred to as figural or negative after-effects and a variety of explanations for them has been proposed. Recently Malhotra reported that similar spatial after-effects occur following the "imaging" of a slanted line,[1] a finding not unlike that of Erikson and Erikson who claimed the occurrence of a negative after-image following the trance induced hallucination of chromatic stimulation.[2] This latter effect, however, was not confirmed by Hibler.[3]

These reports of visual after-effects following imaged or hallucinated stimuli have an important bearing on explanations of these phenomena as well as on experimental methodology and, therefore, warrant more detailed and controlled investigation. Two interpretations of the data can be suggested. First, it can be argued that the judgment made following prolonged imaged stimulation is a consequence of perceptual processes essentially similar to those following actual stimulation. Secondly, in view of recent emphasis on the effects of "demand characteristics" on judgmental responses these data can be interpreted in terms of subject compliance.[4] That is, when the subject (*S*) is motivated to conform to the experimenter's (*E*'s) expectation of the experimental outcome, his responses may be affected accordingly. The occurrence of demand characteristics in an experimental situation can be attributed to *S*'s preconceived notions about the hypotheses involved, to explicit or implicit cues provided by *E*, or to both of these.

---

* Received for publication March 12, 1964. This work was supported in part by the U.S. Public Health Service, N.I.M.H. Grant, M.3950.

[1] M. K. Malhotra, Figural after-effects: An examination of Köhler's theory, *Acta Psychologica, 14,* 1958, 161-198.

[2] M. H. Erikson and E. M. Erikson, The hypnotic induction of hallucinatory color vision followed by pseudo negative after-images *J. exp. Psychol., 22,* 1938, 581-588.

[3] F. W. Hibler, An experimental investigation of negative after-images of hallucinated colors in hypnosis, *ibid., 27,* 1940, 45-57.

[4] M. T. Orne, The nature of hypnosis: Artifact and essence. *J. abnorm. soc. Psychol., 58,* 1959, 277-299.

The aim of the experiment reported here was to examine the effect of instructionally varied demand characteristics on the figural after-effect (*FAE*) from an actual slanted line and one that was imaged after it had been exposed.

Method and Procedure

Experimental design. There were six conditions to which independent groups of *Ss* were assigned and these included three sets of instructions: Positive (*Pos.*), Neutral (*Neut.*), and Negative (*Neg.*) under each of the two conditions of stimulus-presentation. These latter were inspection-figure present (*IFP*) and inspection-figure absent but imaged (*IFA*). The positive instructions were designed to inform *Ss* of the normal expected *FAE* from fixation of a slanted line and the negative instructions to inform him of an "opposite" or "not expected" *FAE*. The neutral instruction excluded any information concerning the direction of a *FAE* from a slanted line.

Apparatus. An aluminum disk 21 in. in diameter was so mounted in a black frame that it could be rotated about its center. The back of the disk was graduated in degrees that the angular orientation of a pattern attached to the front could be read from position of the disk relative to a fixed pointer. The stimulus-pattern consisted of a 4.0 in. × 0.125 in. black line drawn symmetrically about the center of a 21 in. diameter white cardboard disk. Two 0.025-in. diameter black dots, one on either side of the line and in line with its center at a distance of 0.5 in. provided an aid to fixation. This pattern was clipped to the metal disk. A headrest was placed at a distance of 24 in. from the disk so that the visual angle subtended by the line was 9°32′ × 0°18′.

Subjects. The *Ss*, 39 men and 45 women, were drawn from an introductory course in psychology. For the three 'instructions' with inspection-figure present there were 10 *Ss* per group, and for the 'instructions' with inspection-figure imaged there were 18 *Ss* per group. The *Ss* were allocated to a group in the order they reported; each condition being completed before the next was begun.

Procedure. Every *S* underwent a three-phase trial—comprising: a pre-test, 2-min. stimulation, and a post-test. In the pre- and post-tests, *S* was required to adjust the black line to the apparent vertical while fixating its center. Each adjustment was begun with the bar physically vertical. The *FAE* for all conditions was the algebraic difference in degrees between pre- and post-test adjustments. Each *S* underwent four trials, two with the black bar slanted 30° to the right, and two with it slanted 30° to the left. Between successive trials there was a 3-min. rest during which *S* wore black goggles to prevent pattern vision. During actual or imaged stimulation, the order of presentation of the slant of the bar for half the *Ss* was left, right, left, right and this order was counterbalanced for the other half of them.

There were two conditions of stimulation. For the *IFP* condition, *S* was required to fixate the inspection-figure for 120 sec. while it was tilted. For the *IFA* condition, *S* was shown the tilted inspection-figure for 30 sec., after which he was shown a simple block design for a period of 2 min. Following this, *S* was instructed to image the 30° slanted line for 2 min. Following the completion of the four trials, the *Ss* in the *IFA* condition were required to rate the vividness of their imaging using the Bett's rating scale for vividness of imagery.[5]

[5] G. H. Betts, The distribution and functions of mental imagery, *Columbia Univ. Contr. Educ.*, No. 26, 1909, 1-99.

For both the *IFP* and *IFA* conditions, prior instructions were given in order to vary the demand characteristics of the task. Those *Ss* assigned to a 'positive' group were informed of the expected direction of the after-effect, those in the 'negative' group of an unexpected or 'opposite' after-effect, and those in the 'neutral' group were given no instructions regarding the direction of the after-effect.

Since the instructions are of importance in this experiment they are given here in full together with *E*'s demonstration of the procedures required of *S*.

Instructions for IFP condition (positive, negative, and neutral). In this experiment your task will be to move this disk itself, until the black line on it looks vertical, by turning the disk itself. Then I will ask you to gaze at a tilted line and then adjust the disk once more. So I will ask you to do three things adjust the line to the vertical, gaze steadily at a tilted line and adjust the line to the vertical again. All clear?

Put your chin in the chin-rest so that you are looking through the cone. Look directly at the center of the black line and set the disk until the line looks to you as if it is vertical. Try to work speedily. Just adjust the disk till the line looks to you to be vertical. Tell me when you have done this . . . Close your eyes.

[*E* tilts line for fixation-period.]

Open your eyes now and look directly at the center of the disk and the black line. I want you to gaze quite steadily at the center of the black line where the two dots are. Keep very steady and do not shift your gaze during this time. It is very important that you do not let your gaze wander. If you wish to blink, that is all right, but always gaze very steadily right at the center of that black line where the dots are untill I tell you to stop. [120 sec. later] Close your eyes for a second.

[*E* adjusts the test-line to the vertical position.]

Open your eyes and quickly set the disk till the black line looks vertical –just how it *looks* to you. Do this as quickly as you can. . . . Now sit back, put these goggles on and rest for a while.

Instructions for IFA conditions (positive, negative, and neutral). In this experiment your task will be to adjust the disk until this black line looks vertical, by reaching out and turning the disk itself. Then I will show you a tilted line that I wish you to image and I will ask you to adjust the line to the vertical again. So I will ask you to do three things: adjust the line to the vertical, image a tilted line that I will show you beforehand and adjust the line to the vertical again. All clear? Put your chin in the chin-rest that you are looking through the cone and open your eyes. Look directly at the center of the black line and set the disk until the line looks to you as if it is vertical. Try to work speedily; just adjust the disk till the line looks to you to be vertical. Tell me when you have done this . . . Close your eyes.

[*E* tilts line for fixation-period—familiarization.]

Open your eyes. Now I wish you to look at this line. Note the angle of tilt and try to remember it just as it is. Thoroughly familiarize yourself with this tilted line. [30 sec. later *E* shows *S* the set of blocks and gives the following instructions.]

I wish you now to turn toward these blocks on the table behind you and arrange them in different piles according to the shape and color of the figures on them. So arrange them that there is a pile of blue triangles, blue circles, blue squares, and blue diamonds. Just sort the blocks into four different piles and make sure that the blue side is showing for each block.

[If *S* finishes too quickly ask him to arrange the blocks so that the yellow, red or green form is showing.]

Now turn around and face the disk again. Close your eyes. Now with your eyes closed all the time, I wish you deliberately to call to mind and retain an image or mental picture in your mind's eye, as it were, of the tilted black line on the

disk that you saw a few minutes ago. I wish you to deliberately try to evoke a visual mental image of this tilted line set at the center of the white disk. I want you to conjure up this visual image and to focus your image of the line at the center of the disk. Focus your attention steadily at the center of this imaged line on the disk and try to maintain this image of the line tilted to the (left or right) till I ask you to stop imaging. Concentrate on the image of the tilted line that you saw before at the center of the disk. Try to evoke an image as accurately and as vividly as you can of the line you saw before tilted to the (left or right) and try to keep this image steady. [120 sec. later *E* adjusts the test-line to the vertical position.]

Open your eyes. Look at the center of the disk and set the disk as quickly as you can so that the black line looks vertical to you—just how it looks to you. . . ."

Instructions for 'positive' expectancy. When you fixate a line tilted to the left, after a while it appears to straighten up. If you are asked to adjust a vertical line immediately afterwards, it will appear to be tilted in the opposite direction and you tend to straighten it by moving it to the left. Therefore the effect of a line tilted to the left is to make an adjustment to the left. The effect of a line tilted to the right is to make an adjustment to the right. This also applies to imaging a line. I don't wish you to make adjustments to the true vertical. I am asking you to adjust it so that it *appears* vertical to you. Do not follow this information, but make sure not to compensate for this effect.

TABLE I

MEAN ANGULAR DEVIATIONS FROM PRE- TO POST-TEST JUDGMENTS

Stimulus	Demand characteristics		
	neutral	positive	negative
present	+.738	+.944	+.421
absent	+.043	+.298	−.253

Instructions for 'negative' expectancy. When you fixate a line tilted to the left and then a vertical line, the latter looks to you as if it is tilted in the same direction as the line you just fixated. If you are asked to adjust a vertical line immediately afterwards, it will appear to be tilted in the same direction as you tend to straighten it out by moving it to the right. Therefore the effect of a line tilted to the left is to make an adjustment to the right. The effect of a line tilted to the right is to make an adjustment to the left. This also applies to imaging a line. I don't want you to make adjustments to the true vertical I am asking you to adjust it so that it *appears* vertical to you. Do not follow this information, but make sure not to compensate for this effect.

RESULTS

The mean pre-test settings averaged over *S*s for each of the 6 groups did not differ significantly from each other.

Means of angular deviations of pre-test from post-test judgments over the four adjustments are shown in Table I for each experimental condition. They indicate a deviation in the direction expected from the fixation of a tilted line and a deviation in the opposite direction. These means indicate symmetry of size of deviation around the neutral condition. Positive demand characteristics instructions led to an increase in the mean deviation in the expected direction by +0.206 when the stimulus was present and

+0.245 when the stimulus was absent. Negative demand characteristics instructions led to a decrease in the mean deviation in the expected direction by 0.317, when the stimulus was present and 0.296 when the stimulus was absent.

This shows a slightly larger effect for the negative conditions, but since it is the same for both control and experimental groups it should not affect the conclusions with regard to any difference between the two neutral groups.

The main effect of stimulus when present was tested over-all for all demand characteristic conditions. An analysis of variance shows that there is a significant effect for the two conditions of stimulation. The *S*s in the *IFP* condition showed a significantly greater angular displacement between the pre- and post-tests than did the *S*s in the *IFA* condition.

The mean displacement score ($\bar{X} = +0.738$) for the control group with neutral instructions is significantly different from zero in the direction of the expected after-effect, and it also is significantly different from the mean displacement score ($\bar{X} = +0.043$) of the experimental group under neutral instructions. This latter mean is not significantly different from zero.

There were individual differences in imagery ratings. The correlations of imagery ratings with mean displacement scores for each of the three experimental groups were −0.01, −0.24, −0.12, for the neutral, positive, and negative instructions respectively. These correlations indicate no consistent relationship between experimental response and vividness of imagery.

Discussion

The results indicate that in this particular situation there is no evidence for the identity of imaging and perceiving. The interpretation of previous research findings that response to an imaged stimulus is the result of perceptual processes similar to the response when an actual physical stimulus is present, must be rejected. In these earlier studies confirming the identity of imaging and perceiving the assumption was usually made that *S*s were naïve about the perceptual phenomenon being investigated. The methodological scheme used in this experiment was not based on such assumptions about knowledge of the *S* with regard to the expected effect. Three types of instructions stressing different expectations were introduced for both perceptual and imaging groups. Responses to these instructions showed a similar compliance in the direction of the effect stressed by the instructions whether the information given was correct or incorrect. The

systematic effect produced by the control group given neutral instructions was not produced by imaging under the same set of instructions. This difference *cannot* be interpreted as indicating different expectations on the part of the experimental *Ss* since all the experimental groups showed similar patterns of response to the positive and negative instructions as did the control groups given the same instructions.

Within a particular context this experiment has confirmed the hypothesis that imaging is distinct from perceiving when adequate control is exercised over expectancy of the outcome of stimulation.

Summary

The *Ss* (84 in number) were allocated in the order in which they reported to one of six conditions. Three different sets of instructions varied their expectations about the effect of pre-stimulation on the perception of tilt. For each instruction group there were two conditions of stimulus-presentation: fixation-stimulus present and fixation-stimulus absent, but imaged. The aim of the experiment was to investigate the effect of variation in demand characteristics on the figural after-effect under the two conditions of stimulus-presentation. The measure used was the angular deviation of the pre-setting to the vertical from the post-setting to the vertical. Results showed a symmetry of means around the groups that were not instructed to expect an effect in a particular direction. *Ss* imaging the tilted line under this set of instructions showed no appreciable deviation from the vertical in their experimental settings, while *Ss* actually fixating the stimulus showed an appreciable experimental response. The interpretation placed on these findings was that, when *Ss*' expectations are experimentally varied in a comprehensive methodological scheme, imaging is found to be a separate and distinct function from perceiving.

CHANGE IN REACTION-TIME AS A FUNCTION OF KNOWLEDGE OF RESULTS

By RUSSELL M. CHURCH and DAVID S. CAMP, Brown University

The early work of Johanson clearly indicates that speed of reaction may be increased by knowledge of performance,[1] and recent work on vigilance provides further evidence for such a relationship.[2] Typically, an abrupt improvement in reaction-speed with the introduction of the knowledge-of-results has been found. In the present experiment, the permanence of the improvement was studied in a conventional performance vs. learning design.[3]

Method. The *Ss* were 40 students drawn from sections of an introductory course in psychology. *S* sat at a small table on which were mounted a reaction-switch and three event-lights. The reaction-switch was a red push-button which, if depressed a distance of 0.015 in. with a force of about 260 gm., would activate a microswitch. A 7-w. white light, mounted 9 in. behind the reaction-switch, served to indicate anticipatory responses, *i.e.* responses during the ready-signal. A 7-w. green light, mounted 4½ in. to the left and 2 in. in front of the white light, served to indicate a response faster than a criterion. A 7-w. red light, mounted 4½ in. to the right and 2 in. in front of the white light, served to indicate a response slower than a criterion.

The warning signal was a noise of equal energy per cycle, filtered to fall off at 12 db. per octave above 1,000 c.p.s. (Grason-Stadler Model 901A Noise Generator set for 'Speech Noise'). It was about 80 db. above 0.0002 dynes per cm.^{2}, and it was administered through a pair of high-fidelity earphones. Instructions also were delivered through the earphones by a magnetic tape. Stationed in an adjoining room, *E* observed *S* through a one-way vision-mirror and recorded the reaction-times (*RT*) to the nearest 0.001 sec. with an electronic clock (Hunter Klockounter Model 120 A). The presentation of the ready signal and of the event-lights was controlled by relay circuitry located in *E*'s room.

There were two experimental treatments, knowledge of results (*K*) and no-knowledge (*NK*). In both cases, the *Ss* were instructed to push the button as fast as possible at the termination of a ready-signal. In the case of the *K*-treatment, the thirty-third percentile of the latency-distribution was calculated for each previous

* Received for publication March 1, 1963. This research was supported by Grant M-2903 from the National Institute of Mental Health.

[1] A. M. Johanson, Influence of incentive and punishment on reaction-time, *Arch. Psychol.*, 8, 1922 (No. 54), 1-53.

[2] See, for example, P. D. McCormick, Performance in a vigilance task with and without knowledge of results, *Canad. J. Psychol.*, 13, 1959, 68-71.

[3] A critical analysis of this design is provided by K. W. Spence, *Behavior Theory and Conditioning*, 1956, 166-179.

day for a given *S*, and the fastest of these was used as the criterion for that *S* on the next day. Immediately after *S*'s response, the green or the red light was turned on to indicate that the speed was more or less than the criterion. (The light went off at the beginning of the next ready-signal.) In the case of the *NK*-treatment, the lights were simply not used. In the case of both treatments, a response during the ready-signal turned on the white light, which stayed on until the beginning of the next ready-signal.

All the *Ss* had the *NK*-treatment on the first day. Half the *Ss* had the *K*-treatment and half the *NK*-treatment on Days 2–4. On Day 5, half the *K-Ss* were shifted to *NK* and half the *NK-Ss* were shifted to *K*.

For the first 20 *Ss*, who were randomly assigned in equal numbers to the four groups, the warning signal was 2-sec. in duration. For the next 20 *Ss*, also randomly assigned in equal numbers to the four groups, the duration of the warning signal varied in 0.1-sec. steps from 1.5–2.5 sec. with a mean of 2.0 sec. The intertrial interval was 10 sec. Each session lasted about 50 min. and consisted of three blocks of trials (83, 84, and 83 trials, respectively). There were rest-periods of 110 sec. after the first and second blocks of trials.

Results. The reaction-time (*RT*) on each trial in units of 0.01 sec. was transformed to log (RT −0.10), 0.10 sec. being considered an approximation to the irreducible minimum for this situation.[4] (Any response faster than 0.10 sec. was considered an 'anticipatory reaction,' and it was discarded from the latency-analysis.) On the first day of practice, the antilog of the mean of the transformed *RT*s of the *Ss* with the constant warning signal was 208 m.sec. as compared to 240 m.sec. for *Ss* with the variable signal ($t = 6.4, p < 0.01$).

To reduce the dependence of subsequent performance on the initial level, the transformed *RT*s for subsequent days were expressed as a percentage of the transformed *RT* for the first day.

The performance of the four groups as a function of days of practice is shown in Fig. 1. An analysis of variance (Lindquist's Type III) for the data of Days 2–4 showed significant effects of the three main troubles.[5] The *Ss* with the fixed warning signal were significantly faster than those with the warning signal of variable duration ($F = 33.3$, $df = 1, 36$, $p < 0.01$); *Ss* with knowledge of results were significantly faster than those without knowledge of results ($F = 14.3$, $df = 1{,}36$, $p < 0.01$); and the *RT* decreased over days ($F = 31.4$, $df = 2{,}72$, $p < 0.01$). The practice-effect was greater in the case of *Ss* with the fixed warning signal than in the case of those with the warning signal of variable duration ($F = 6.6$, $df = 2{,}72$, $p < 0.01$), and greater for *Ss* with knowledge than without

[4] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, 1954, 21.
[5] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1953, 281-284.

($F = 4.7$, $df = 2{,}72$, $p < 0.05$). The remaining two interactions did not approach significance. If the differential treatment during the 750 trials of Days 2–4 had any lasting effects, the performance on Day 5 should reflect that treatment. The mean transformed *RT* on Day 5 is shown in Table I as a function of the treatment on Days 2–4 and on Day 5 for *Ss* with fixed and variable signals. A three-way analysis of variance showed that *Ss* with the fixed warning signal had a greater reduction in *RT* than did those with a variable signal ($F = 42.8$, $df = 1{,}32$, $p < 0.01$), and that *Ss* with knowledge on Day 5 responded more quickly than those with-

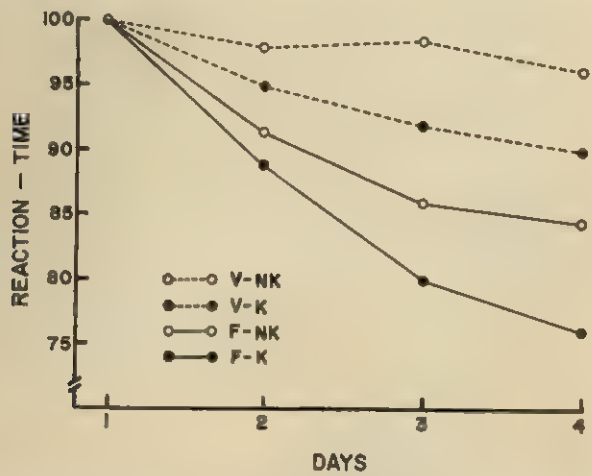

FIG. 1. MEAN TRANSFORMED REACTION-TIME ON DAYS 1–4

Expressed as a percentage of the mean transformed *RT* of Day 1, for *Ss* with warning signal of fixed (*F*) and of variable (*V*) duration under conditions of knowledge of results (*K*) and no knowledge (*NK*).

out knowledge ($F = 10.3$, $df = 1{,}32$, $p < 0.01$). Despite the suggestion in Table 1 of the relevance of the treatment during Days 2–4 for the performance on Day 5, the analysis of variance did not demonstrate the significance of this factor ($F = 2.7$, $df = 1{,}32$, $p > 0.05$).

The transformed *RT*s expressed as a percentage of Day 1 remained fairly constant for all groups on Day 5. Only in the case of the group with a fixed warning signal and no knowledge on Day 5 did the effects of treatment on Days 2–4 tend to disappear gradually on Day 5. In the other three comparisons, the effect of the treatment on Days 2–4 was no more apparent on the first block of trials of Day 5 than on the last block.

Discussion. If information about speed of reaction serves as reinforcement, then knowledge of results provides differential reinforcement of short-latency responses. Such differential reinforcement should result in (*a*) gradual improvement during the treatment, (*b*) gradual decrement after cessation of the treatment; and (*c*) a persisting influence on later trials on which *S*s are treated alike. In studies of *RT* in a competitive situation—the *S*s being told which of two competitors was faster on each trial—the *RT*s are reliably shorter than those of control *S*s under normal conditions.[6] The effect of 50 such trials could not, however, be detected when

TABLE I

MEAN TRANSFORMED REACTION-TIME ON DAY 5

		Treatment on Day 5						
		Fixed warning signal				Variable warning signal		
		NK	*K*	mean		*NK*	*K*	mean
Treatment on Days 2–4	*NK*	0.864	0.739	0.801	*NK*	0.959	0.926	0.942
	K	0.796	0.717	0.757	*K*	0.938	0.886	0.912
	Mean	0.830	0.728		Mean	0.948	0.906	

*S*s were returned to the non-competitive condition. In the present experiment, the effect of 750 trials of training under *K* or *NK* conditions failed to produce a difference detectable on 250 subsequent trials in which *S*s were treated alike. During the extended series of trials, there was a notable practice-effect, and the interaction between trials and treatment suggests that the practice-effect may have been greater for the *S*s with knowledge of results than for those without that knowledge. This result should not be overemphasized, however, since the divergent curves could be made parallel or convergent by the use of some other measure of the dependent variable.

Although no formal measures of level of motivation were made in this experiment, the *S*s without knowledge typically seemed far more bored than those with knowledge. Most of the latter *S*s appeared more alert and several were highly activated. One *S*, to cite an extreme example, often would shout and pound the table when the red light appeared. Although it is probable that knowledge of results increased the level of motivation, it is not at all clear that the increase was causally related to the increase in speed of performance.

[6] R. M. Church, The effects of competition on reaction-time and palmar skin conductance, *J. abnorm. soc. Psychol.*, 65, 1962, 32-40; R. M. Church, R. B. Millward and P. Miller, Prediction of success in a competitive reaction-time situation, *ibid.*, 1963, 67, 234-240.

Summary.—The purpose of this experiment was to study the permanence of the decrease in *RT* produced by knowledge of results. Forty *S*s were given 1250 *RT*-trials in a period over five days. The results were: (a) *S*s with knowledge were faster than those without knowledge; (b) *S*s with a warning signal of fixed duration were faster than those with a warning signal of variable duration; and (c) the *RT*s decreased as a function of practice. There was no evidence of any lasting effect of knowledge on subsequent performance. Knowledge of results seemed to be effective only during periods in which it was given.

TASTE-SENSITIVITY AFTER EATING: A SIGNAL-DETECTION APPROACH

By MARY E. MOORE, EUGENE LINKER, and MARGUERITE PURCELL, *
University of Pennsylvania

The recent development of a signal-detection model based on statistical decision-theory has called into question much of traditional psychophysics. A distinction has been made between *S*'s sensitivity and his response-bias, and the whole notion of an absolute sensory threshold has been challenged.[1] Previous work from this laboratory has demonstrated that detection-methodology, developed largely in the areas of audition and vision, could be adapted for the study of taste and has yielded results showing that the simple threshold-model also is inappropriate in the gustatory mode.[2] The present study illustrates the application of the detection-model to a specific problem in the area of taste.

Several previous investigators, using some variation of the method of limits, have reported an increase in taste-threshold following eating, which has been attributed to a decrease in need. Goetzl, Ahokas, and Payne reported a rise in threshold following lunch as compared with no-lunch,[3] and similar results with sodium chloride subsequently were reported by Irvin and Goetzl.[4] Hammer studied the effect of lunch on sensitivity to hydrochloric acid.[5] His data revealed a pattern like that found in the previous studies and a decrease in sensitivity following lunch. Yensen investigated taste-thresholds for several substances in two *Ss* under a variety of conditions of food-intake.[6] He found that thresholds were raised following breakfast as compared with the fasting state, and a large lunch produced a rise in threshold as compared with the normal and no-lunch conditions. Studies by Furcht-

* Received for publication July 17, 1963. This research was supported in part by Grant M-3684 from the National Institute of Mental Health, U.S. Public Health Service. We thank R. Duncan Luce, Eugene Galanter, and Albert Stunkard for their criticism and advice.

[1] John Swets, Wilson Tanner, and Theodore Birdsall, Decision processes in perception, *Psychol. Rev.*, 68, 1961, 301-304; David Green, Psychoacoustics and detection theory. *J. acoust. Soc. Amer.*, 32, 1960, 1189-1203.

[2] Eugene Linker, M. E. Moore, and Eugene Galanter, Taste thresholds, detection models and disparate results, *J. exp. Psychol.*, 67, 1964, 59-66.

[3] Franz Goetzl, Ann Ahokas, and Jean Payne, Occurrence in normal individuals of diurnal variations in acuity in the sense of taste for sucrose, *J. appl. Physiol.*, 2, 1950, 619-626.

[4] Dona Irvin and Franz Goetzl, Diurnal variations in acuity of taste for sodium chloride, *Proc. Soc. exp. Biol. Med.*, 79, 1952, 115-118.

[5] F. J. Hammer, The relation of odor, taste, and flicker-fusion thresholds to food intake, *J. comp. physiol. Psychol.*, 44, 1951, 403-411.

[6] Roy Yensen, Some factors affecting taste sensitivity in man: Food intake and time of day, *Quart. J. exp. Psychol.*, 11, 1959, 211-229.

gott and Friedman[7] and by Pangborn[8] both produced some slight evidence that thresholds were raised after eating, but in both cases the rise was not significant. Contradictory evidence has been reported by other investigators employing the traditional methodology. Janowitz and Grossman, studying thresholds at 10:30 A.M., and before and after a noon meal, concluded that there was no relationship between eating and the ability to taste sucrose or sodium chloride.[9] Meyer found no systematic changes in threshold for sodium chloride, quinine sulfate, or glucose during a period of 34 hr. of food-deprivation.[10]

The present experiment, designed to study the effect of a noon meal on sensitivity to sucrose, represents a critical departure from the previous work. In the first place, we employed a *"yes-no"* procedure in which only one stimulus-concentration, *i.e.* one signal-strength, was presented to *S*. Moreover, the test-solution or 'signal' was presented at only half of the trials according to a random schedule; at the remaining half there was only 'noise' *i.e.* distilled water presented to him.

Procedure. From preliminary tests, a solution of sucrose was selected for each *O* that he could distinguish correctly from water only about 70-80% of the time.[11] In other words, the signal-strengths selected were well within the range of uncertainty. They varied slightly from one *S* to another, ranging from 0.200–0.275%, *i.e.* from $5.8–8.0 \times 10^{-3}$ M.

Seven *Ss*, three men and four women, were tested in two 50-min. sessions per day for 10 successive days. The first session was at 11 A.M. and the second at 1 P.M. The *Os* refrained from eating or smoking for 11 hr. prior to each morning session. On 5 of the 10 days, selected according to a random schedule, the *Ss* ate a noon meal following the morning session. They were not told until after the morning session for a particular day whether they were to eat lunch or to fast.

On each of the 70 trials within each session, *S*, blindfolded, was presented with a small glass containing 1 ml. of the test-liquid. After tasting each sample, *S* reported whether he believed it to have been 'sugar' or 'water,' and was immediately informed as to the correctness of his response by means of a tone delivered through earphones. The electronic signal eliminated conversation between *E* and *S*, and the earphones tended to cut down on distracting noises which interfered with *S*'s, concentration. A 30-sec. intertrial interval followed the feedback-signal. The end of this interval and the beginning of the next trial were indicated by another electronic signal.

To maximize motivation, the *Ss* were rewarded for their participation in the experiment solely on the basis of the accuracy of their performance. For each correct

[7] Ernest Furchtgott and M. P. Friedman, The effects of hunger on taste and odor RLs, *J. comp. physiol. Psychol.*, 53, 1960, 576-581.

[8] R. M. Pangborn, Influence of color on the discrimination of sweetness, this JOURNAL, 73, 1960, 229-238.

[9] H. D. Janowitz and M. I. Grossman, Gusto-olfactory thresholds in relation to hunger and appetite, *J. appl. Physiol.*, 2, 1949, 217-222.

[10] D. R. Meyer, The stability of human gustatory sensitivity during changes in time of food deprivation, *J. comp. physiol. Psychol.*, 45, 1952, 373-376.

[11] Much of the procedure has been described in detail elsewhere and will not be repeated here. See Linker, Moore, and Galanter, *op. cit.*, 59-66.

response (either saying "sugar" when the stimulus was a sucrose solution, or saying "water" when it was distilled water), *S* received five cents; for each incorrect response five cents were deducted from his earnings.

Results. The results were tabulated in terms of the percentage of 'hits' and 'false alarms' made by each *S* for each session. A 'hit' is defined as a response of 'sugar' when sugar was in fact presented; a 'false alarm' is a response of 'sugar' when water was presented. These data then were transformed into another pair of numbers according to the following theoretical model: It is assumed that any taste-sensation experienced by *S* in the experiment can result either from signal-plus-noise (sucrose plus distilled water) or from noise alone (plain distilled water). It is further assumed that the likelihood of any given sensation's arising from noise alone or from a signal-plus-noise is described by two overlapping probability-density curves. When confronted with any of the infinite number of sensations lying along a continuum, *S* evaluates it in terms of a decision-line or criterion (c'). If the sensation exceeds the criterion, he reacts as if it were a signal; if it is below, he reacts as if it were merely noise. Thus the area of hits is represented by the area under the curve generated by the signal-plus-noise probability-distribution that lies to the right of the criterion. Similarly, the area under the curve generated by the noise-distribution that lies to the right of the criterion represents the area of false alarms. The separation between the curves (d') can thus be seen to constitute a measure of sensitivity since the larger this distance, the less the overlap between the false alarms and hits, and the greater the sensitivity. On the assumption that the signal-plus-noise and noise-distributions both are normal, this distance (d') can be expressed simply in terms of the normal deviates corresponding to the hit and false-alarm probabilities, which in turn can be estimated empirically from the percentage of hits and false alarms made by *S*. In calculating d', the assumption commonly made is that the *SD*s of the signal-plus-noise and the noise-distributions are equal, *i.e.* that their ratio is 1. It is preferable, however, to use an empirical estimate of this ratio whenever possible.

For the three men studied, empirical estimates of the ratio of the *SD* of the signal-plus-noise distribution to the *SD* of the noise-distribution were available from previous work and were used in calculating d' values.[12] These ratios were: 1.69, 1.04, and 1.40. For the four women studied, no empirical data were available, and the statistical evaluation of

[12] Iso-sensitivity or so-called ROC curves had been obtained for these *O*s and the ratios were estimated from the reciprocals of the slopes of these curves on a normal normal plot (Linker, Moore, and Galanter, *op. cit.*, 59-66).

their data was based on the very conservative assumption that the true ratios lay somewhere in the range of 1.00-2.00. We calculated for each woman, not one, but a set of d'-values corresponding to this range. For the statistical tests, we selected the values which provided the most stringent test of each null hypothesis.

In contrast to the earlier reports of decreased sensitivity following eating, the present data exhibited a slight increase in sensitivity from A.M. to P.M. in the Lunch as well as the No-lunch conditions. A t'-test revealed no statistical difference between the two conditions in this respect, and when the Lunch and No-lunch data are combined, the over-all A.M. to P.M. increase in sensitivity was significant ($t = 3.39$, $p < 0.02$). Eating *per se* did not, then, appear to affect sensitivity to sucrose, although there was a general increase in sensitivity from morning to afternoon. Previous investigators have noted a similar increase in fasting *Ss*, which may represent a diurnal variation in the ability to taste, or simply an intra-day practice-effect. Since we failed to find evidence for inter-day practice effects in these well-practiced *Ss*, the former explanation seems more likely.

An estimate of the criterion, c', was obtained from the normal deviate corresponding to *S*'s false-alarm rate. Following lunch, *Ss* reported "sugar" less often, and this shift in criterion was statistically significant ($t = 3.63$, $p < 0.012$). No such shift was observed following the No-lunch period, four of the seven *Ss* actually shifting their criteria in the opposite direction.

Although the difference between the Lunch and No-lunch conditions is not statistically significant, the significant shift following lunch is worth noting because it could account for the decrease in sensitivity reported by some of the earlier investigators.

The performance of an *S* in a method-of-limits experiment is, to use the language of signal-detection, based entirely on his ability to make hits without taking into account his false alarms. In the absence of systematic information on the false-alarm rate, a change in the *S*'s so-called 'threshold' that results from an actual change in the sensitivity cannot be distinguished from a similar change in 'threshold' that results from a change in criterion with no change in sensitivity. Had the *Ss* in the previous experiments changed their criteria after lunch in the same way in which ours did, they would have responded with "sugar" less often and would therefore have made fewer correct responses, since in the method-of-limits-design sugar was presented on every trial. Thus, a mere shift in criterion would explain the earlier findings of an apparent decrease in sensitivity after eating. Until a sensitivity-change due to eating has been demonstrated in an experi-

mental design which allows for an estimation of the S's criterion, the negative results of the present experiment must take priority.

Summary. A signal-detection design which makes the distinction between sensitivity and response-bias was used to study the effect of eating on the ability of Ss to taste a sucrose solution before and after a lunch and a no-lunch condition. Sensitivity was found to be higher in the afternoon as compared with the morning sessions, but there was no difference in the sensitivity when the lunch and no-lunch conditions were compared. There was, however, a significant bias on the part of the S to report tasting sugar less after eating. Results of previous studies using the method of limits may have been confounded by the effects of similar bias.

AN EXPERIMENTAL COMPARISON OF THE METHOD OF LIMITS AND THE DOUBLE STAIRCASE-METHOD

By Jacob Nachmias and Robert M. Steinman,
University of Pennsylvania

It is a well-known psychophysical fact that the values of thresholds and *PSE*s depend in part upon the method of measurement.[1] In the method of constant stimuli, the distribution of values of the stimulus-variable presented to *O* affects at least measures of central tendency. In the traditional method of limits, the end-point of a run is undoubtedly biased by its starting point. This bias may be particularly serious when only ascending or only descending runs are employed, and when the variability of the end-points is a function of the parameter under investigation. It may even be present, however, when the two kinds of runs are interspersed, because the amount of the bias probably depends on the distance separating the starting point and the 'true' threshold or *PSE*. Thus, the method of limits permits *E* to influence his results, even if only inadvertently.

Cornsweet has recently described a variant of the staircase-method (itself a variant of the method of limits) which appears to be largely free of starting point bias.[2] In the usual staircase-method, or method of "up and down,"[3] *E* changes the value of the independent stimulus-variable by the same amount on every trial, but in a direction completely determined by *O*'s response on the just preceding trial: The direction of stimulus-change is reversed when *O*'s response changes, but not otherwise. In the new version, trials from two such staircase-series with different starting points are randomly mixed. This paper presents data from an experiment on the detection of a line-increment, designed to compare Cornsweet's double stair-

* Received for publication August 28, 1963. This investigation was supported by research grant B-3682(C1) from the National Institute of Neurological Diseases and Blindness, Public Health Service.

[1] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, 1954, 225-233; W. S. Verplanck and J. W. Cotton, The dependence of frequencies of seeing on procedural variables: I. Direction and length of series of intensity-ordered stimuli, *J. gen. Psychol.*, 53, 1955, 37-47.

[2] T. N. Cornsweet, The staircase-method in psychophysics, this Journal, 75, 1962, 485-491.

[3] W. J. Dixon and F. J. Massey, *Introduction to Statistical Analysis*, 1957, 279-286.

case-method and the traditional method of limits with regard to bias introduced by the starting point.

Experimental conditions. O was dark-adapted for at least 15 min., after which he looked at a small red fixation-point, while steadying his head with the help of a dental impression-plate. The stimulus-object consisted of a bright vertical line 2′ of arc wide superimposed on a disk of white light 1° in diameter and located 10′ of arc above the fixation-point. The disk and line appeared together for 52.4 m.sec. every 10 sec.; a warning buzzer preceded each trial by 2 sec. The luminance of the disk was fixed at 0.42 log m.L. while that of the line increment was varied in 1.1-db. steps in accordance with the particular psychophysical method employed. The details of the optical system, shutter-system, and method of calibration have all been reported previously.[4]

Observers. Both *O*s were men with at least 20:20 vision, uncorrected. One of them (*RS*) was the junior author, who knew well both the psychophysical methods employed and the purpose of this investigation. He was told, however, neither how many different starting points were to be tested nor the starting point of any particular series. The other *O* (*VV*) was told nothing about the methods or aims of the experiment, nor did he discover what they were during the course of the experiment.

Procedures. Each experimental session consisted either of 8 or of 4 double-staircase series and a like number of pairs of limits. The methods were alternated within each session, beginning with double staircases on half of the sessions. As shown in Table I, four different pairs of starting luminances of the line-increment were employed with both methods: two were high bias (*H* and *H′*) and two were low bias (*L* and *L′*). All four pairs were tested equally often in each session, but in a different, scrambled order. Furthermore, the lower of each pair of starting luminances was shown first randomly half the time.

The reference-level around which the starting points were centered was determined separately for each *O* on the basis of his performance in a few preliminary sessions. In fact, *RS*'s reference-level turned out to be 2.4 db. lower than *VV*'s (—0.09 log m.L. vs 0.15 log m.L.). After the reference-level was determined, both *O*s participated in several experimental sessions. The data reported are based on the last seven presentations of the four bias-conditions with each of the methods.

Every run by the method of limits terminated as soon as *O* changed his response. Every double staircase lasted for 32 trials. These trials were not random mixtures of two staircase-series, as Cornsweet recommends.[5] Rather, random permutations of the numbers 1 to 16 were used in such a way as to insure that the two staircases were represented by equal numbers of trials in the first and second halves of each double staircase-series.[6] This procedure made it possible to compare *O*'s performance on the first and second 16 trials without having to take account of the relative number of trials coming from the two staircases.

Results and discussion. The end-point of a run by the method of limits

[4] Jacob Nachmias and R. M. Steinman, Study of absolute visual detection by the rating scale method, *J. opt. Soc. Amer.*, 53, 1963, 1206-1213.

[5] Cornsweet, *op. cit.*, 489-491.

[6] W. G. Cochran and G. M. Cox, *Experimental Designs*, 2nd ed., 1957, 583-595.

was taken to be the decibel-mean of the luminance of the trial on which *O*'s response changed and the luminance on the just preceding trial. Similar means were obtained in double staircase-series, except that the just preceding trial from the appropriate staircase was considered, even if several trials intervened from the other staircase. For every occasion on which each of the four pairs of starting points was used, the means of the following end-points were computed: *double staircases*—(a) the first from each staircase; (b) the last from each staircase; (c) all of those occurring between Trials 1 and 16; (d) all of those occurring between Trials 16 and

TABLE I

Starting Luminances and Grand Means of End-Point Luminances, Expressed in Decibels re Reference-Level

Methods	*O*	Starting points: high bias *H*	*H'*	low bias *L*	*L'*
		5.10–1.67	4.00–0.55	1.75–4.88	0.57–3.80
Limits	*RS*	2.10	2.78	1.29	1.38
	VV	2.42	2.25	1.56	2.05
Staircase:					
First two	*RS*	2.45	2.19	0.00	1.01
	VV	2.40	2.63	2.21	0.84
Last two	*RS*	2.28	1.26	2.10	2.03
	VV	2.35	2.34	2.28	2.80
Trials 1–16	*RS*	2.62	2.04	1.01	1.10
	VV	2.52	2.65	2.19	1.24
Trials 17–32	*RS*	1.85	2.04	1.86	2.02
	VV	2.79	2.43	2.33	2.68

32; *limits*—the two end-points. Table I contains the grand means of these five measures.

The starting point clearly affects the average end-point in the method of limits, in agreement with the previous findings of Verplanck and Cotton.[7] An analysis of variance of these data revealed a significant effect of starting point ($F = 7.31$, $df = 1/52$, $P < 0.01$). Both *O*s were affected in much the same way: the *O* by starting point interaction was not statistically significant. The error-variance was 1.49 db.² In the double staircase-method, the effect is, if anything, more pronounced, if one considers only the end-points within the first 16 trials. Separate analyses of

[7] Verplanck and Cotton, *op. cit.*, 37-42.

variance on the end-points from the first and last 16 trials were performed. For these analyses, the two pairs of high bias starting points (H and H') were considered together, as were L and L'. The effect of the starting points on the end-points was statistically significant only in the first 16 trials ($F = 13.75$, $df = 1/52$, $P < 0.001$). The O by starting point interaction was not significant in either analysis. The error-variance in the analyses of the first 16 trials was 1.21 db.^{2} and in the last 16 trials, 0.97 db.^{2} The magnitude of the bias is not very large (about 1.0 db.), but it constitutes 25% of the decibel-difference between the mean starting points of the H-H' series and the L-L' series. Larger differences between mean starting points possibly would have produced larger effects.

On the other hand, two other measures from the double staircase-method appear to be unaffected by the locations of the starting point; namely, the last end-points and the mean end-point in the last 16 trials. These same measures, however, may not escape contamination if mean differences between starting points larger than 3.3 db. are tested. Since this psychophysical method will generally be used to find the value of an unknown threshold or *PSE*, it may be necessary to proceed beyond 32 trials before end-points reach a steady level unaffected by the starting points.

TRANSPOSITION AFTER 'SINGLE-STIMULUS' TRAINING

By N. J. Mackintosh, Oxford University, England

In a series of experiments, Lashley and Wade trained rats, first to discriminate between a circle and a black card, and secondly to discriminate between the original circle and a new one differing from it in size.[1] Finding no evidence that *S*s would continue to select the original circle if it had been reinforced, or continue to avoid it if it had not been reinforced, they claimed that gradients of generalization were not formed after training with a single stimulus. Grice repeated Lashley and Wade's two experiments in a different situation and found results more in accord with Hullian notions of generalization.[2] Grice's results are, in fact, frequently cited as providing decisive disproof of Lashley and Wade's theory.[3] The present experiments demonstrate that, whatever may be the merits of Lashley and Wade's theoretical position, Grice's two experiments do not provide an appropriate test of it.

A close inspection of Grice's results reveals two surprising facts. (1) In his second experiment, the final size-discrimination was learned by the nonreversal group with 3.7 errors, and by the reversal group with 11.7 errors, while naïve *S*s learn the same discrimination in the same apparatus with 14.1 errors—*i.e.* with more errors than *S*s supposedly learning a reversal.[4] (2) In Grice's first experiment, the reversal group scored above chance over Trials 1-20 of the reversal. To explain this, Grice suggested that rats have a marked preëxisting preference for the larger of two stimuli in a size-discrimination (*S*s were being trained with the large circle positive). In Grice's second experiment, all Ss were trained with the small circle positive and there, too, the reversal group scored above chance over

* This research was supported in part by the U.S. Office of Naval Research under terms of Contract N62558-3927, and in part by the Nuffield Foundation.

[1] K. S. Lashley and Marjorie Wade, The Pavlovian theory of generalization, *Psychol. Rev.*, 53, 1946, 72-87.

[2] G. R. Grice, The acquisition of a visual discrimination habit following response to a single stimulus, *J. exp. Psychol.*, 38, 1948, 633-642; The acquisition of a visual discrimination habit following extinction of response to one stimulus, *J. comp. physiol. Psychol.*, 44, 1951, 149-153.

[3] C. E. Osgood, *Method and Theory in Experimental Psychology*, 1953, 449; G. A. Kimble, *Hilgard and Marquis' Conditioning and Learning*, 1961, 369-371.

[4] Grice, Visual discrimination learning with simultaneous and successive presentation of stimuli, *J. comp. physiol. Psychol.*, 42, 1949, 365-373.

Trials 1-20. It can hardly be supposed that rats prefer both large to small, and small to large.

These points seem to imply that the "reversal" groups in Grice's experiments were not learning a reversal, and that Grice's analysis of the situation is incorrect. An alternative analysis is that in the first stage of his experiments, where *S*s were discriminating a circle from a black door, they were in fact learning to approach the *brighter* of two stimuli (if the circle was positive), or the *darker* of two stimuli (if the circle was negative). Since a large white circle is brighter than a small white circle, the second stage of the experiments, in which *S*s were discriminating two circles of different size, also was soluble as a brightness-discrimination. Now, neither

TABLE I

THE DESIGN OF GRICE'S TWO EXPERIMENTS
(0 = black door; 5 = 5-cm. circle; 8 = 8-cm. circle)

Experiment	Group	Stage 1		Stage 2	
		+	−	+	−
I	−−	0	8	5	8
	−+	0	5	5	8
II	++	8	0	8	5
	+−	5	0	8	5

Lashley and Wade nor Grice used a factorial design in their experiments. The design of their experiments is symbolized in Table I, from which it can be seen that although, for the reversal groups, the sign of the circle shown in Stage 1 was reversed in Stage 2, no *S*s were required to reverse the brightness-discrimination established in Stage 1: in Experiment I, *all* *S*s were required to select the darker (smaller) of two stimuli; in Experiment II, *all* *S*s were required to select the brighter (larger) of two stimuli. If this analysis is correct, then to obtain a reversal group it would be necessary to train *S*s in Stage 2 in the opposite direction to that used in these earlier experiments.

Method. The *S*s were 32 naïve female hooded rats, 2.5 mo. old. They were trained in a modified Lashley jumping stand, with landing platforms below each stimulus-window. The stimuli were circles, 5 and 8 cm. in diameter, respectively, cut from ¼-in. white Perspex, and mounted on brown hardboard doors. The third stimulus was a blank hardboard door. The *S*s were accustomed to a 24-hr. feeding schedule, habituated to the apparatus, and trained to jump gradually increasing distances through open doors. During training, 20 noncorrection trials were given daily, with a 6–8 min. intertrial interval. No punishment was given for incorrect response. After each correct response, *S*s were allowed to eat for 20 sec. The cri-

terion of learning was 18 correct responses out of any 20 consecutive trials with the last 10 all correct.[5]

Experimental design. The experiment consisted of two stages. In Stage 1, 16 *Ss* were trained to jump to a blank hardboard door and away from a circle, and the remaining 16 were trained with a circle positive and the blank-door negative. Within each group, 8 *Ss* were trained with the 5-cm. circle, and 8 with the 8-cm. circle. Upon reaching criterion, all *Ss* were given 20 additional trials, and then began Stage 2. In Stage 2, all *Ss* learned to discriminate between the two circles. Half of each of the above subgroups were trained with the 5-cm. circle positive, half with the 8-cm. circle positive. There were, therefore, 8 groups (each of 4 *Ss*) trained as shown in Table II.

Results. Both on Grice's and on the present analysis, Groups A-I and B-I learn a nonreversal in Stage 2, while groups A-II and B-II learned a reversal. The critical comparison is between Groups III and Groups IV: According to Grice, Groups A-III and B-III must learn a reversal, Groups A-IV and B-IV a nonreversal. If, on the other hand, *Ss* learn a brightness-discrimination in Stage 1, and transfer it to Stage 2, then Groups A-III and B-III can continue to respond in the direction in which they were initially trained, while Groups A-IV and B-IV must learn a reversal.

TABLE II

EXPERIMENTAL DESIGN AND RESULTS

The first two rows show the positive and negative stimuli for each group in Stages 1 and 2. 0 = blank door; 5 = 5-cm. circle; 8 = 8-cm. circle; the positive stimulus is always shown first.

	Groups							
	A-I	A-II	A-III	A-IV	B-I	B-II	B-III	B-IV
Stage-1 stimuli	0–8	0–8	0–5	0–5	8–0	8–0	5–0	5–0
Stage-2 stimuli	5–8	8–5	5–8	8–5	8–5	5–8	8–5	5–8
Percent correct (Trials 1–20)	75.0	26.5	70.0	37.5	56.2	35.0	73.7	32.5
Trials to criterion	16.0	185.0	30.2	160.5	46.0	137.2	53.5	137.5

The main results of the experiment are shown in Table II. Over Trials 1-20, Groups A-IV and B-IV scored below chance, while Groups A-III and B-III scored above chance, and an analysis of variance showed this difference to be highly significant ($F = 44.73$, $df = 1/13$, $p < 0.001$). The trials-to-criterion scores show even larger differences: Group A-IV averaged rather more than five times as many trials as Group A-III (the fastest *S* in A-IV learned in 142 trials; the slowest in A-III learned in 39); Group B-IV averaged nearly three times as many trials as Group B-III (the fastest

[5] Further details of apparatus and training procedure are given in N. J. Mackintosh, The effect of irrelevant cues on reversal learning in the rat, *Brit. J. Psychol.*, 54, 1963, 127-134.

S in B-IV learned in 122 trials; the slowest in B-III learned in 73). There can, therefore, be no doubt whatsoever that Grice's analysis is incorrect, and that it is Groups IV, not Groups III, for which Stage 2 is a reversal.

It remains true that, in terms of trials to criterion (but not in terms of percentage correct over Trials 1-20), the differences between Groups I and Groups III are comparable to the differences obtained by Grice in his two experiments—although in the present experiment, owing to the small number of *S*s, they fell short of significance ($F = 3.73$, $df = 1/12$, $0.10 > p > 0.05$). This result implies that *S*s had learned something about the absolute properties of the circles used in Stage-1 training, and that a small absolute effect is superimposed on a much larger relational effect. This is irrelevant, however, to the main contention of the present paper, which is that the experiments of Lashley and Wade and of Grice do not provide a test of generalization following single-stimulus training, because they did not involve *single*-stimulus training at all.

It is not claimed that these results resolve the theoretical disagreement between Lashley-Wade and Grice; they serve rather to set the record straight. Although it seems clear that single-stimulus training can produce significant gradients of generalization,[6] it has recently been suggested that those experiments which have produced results at variance with Lashley and Wade's position have in fact involved some discriminative training.[7] It also is clear that part of Lashley and Wade's claim receives strong support from experiments demonstrating that the nature of prior discriminative training can significantly affect the slopes of generalization-gradients.[8]

Summary. Lashley-Wade's and Grice's experiments, in which rats were pretrained to jump to or away from a single circle and then tested on a size-discrimination, were repeated. It was shown that such pretraining involves the learning of a brightness-discrimination between *two* stimuli—the dark, blank door, and the brighter circle—and that Grice's results therefore are not critical for Lashley and Wade's theoretical position.

[6] An experiment by Spence quoted in C. L. Hull, The problem of primary stimulus stimulus generalization, *Psychol. Rev.*, 54, 1947, 120-134.

[7] E. G. Heinemann and R. L. Rudolph, The effect of discriminative training on the gradient of stimulus-generalization, this JOURNAL, 76, 1963, 653-658.

[8] E. F. MacCaslin, Jerome Wodinsky, and M. E. Bitterman, Stimulus-generalization as a function of prior training, this JOURNAL, 65, 1952, 1-15; Mackintosh, The effect of attention on the slope of generalization gradients, *Brit. J. Psychol.*, 56, 1965, 87-93.

NONRELATIONAL JUDGMENTS OF SIZE AND DISTANCE

By WILLIAM EPSTEIN, University of Kansas

Experiments which purport to demonstrate the influence of specified variables on the perception of *absolute* size or distance[1] are frequently reinterpreted as dealing with relational effects.[2] It also has been asserted that any experiment employing a visual comparison-object deals, of necessity, with relational effects.[3] These arguments create uncertainty about a recent demonstration that 'assumed size' can determine absolute perceived distance.[4] It was decided, therefore, to repeat the original assumed-size experiment employing nonvisual (tactual) measures of the size and distance of individual 'coins' presented usually without any accompanying visual stimuli. If visual relationships were responsible for the original results, then the elimination of these relationships should give different results.

Apparatus. Each standard was mounted on a thin black rod and placed on a black table-top at approximately eye-level. Provisions were made to restrict *O* to stationary, monocular, aperture-vision when desired. It was also possible to present the standard in isolation in surrounding darkness, illuminated by a spot of light which corresponded exactly to the dimensions of the standard.

Materials. Two sets of standards were used. The training set consisted of three copper disks with diameters of 1.75, 2.38, and 2.86 cm., respectively. The test-standards were the mounted photographs of the dime, quarter and half-dollar which were used in the earlier study. Each of these standards had a diameter of 2.38 cm. —the normal size of a quarter.

The size-comparison series was a set of seven disks cut from one sheet of copper. The disks varied in diameter from 1.43-3.33 cm. in steps of 0.125 cm. They were glued to the surface of a wooden panel, arranged horizontally in order of size. The distance-comparison object was a coiled rope which could be pulled by

* Received for publication April 22, 1964. This study was supported by Grant MH 4153-04 from the National Institute of Mental Health.

[1] W. H. Ittelson, Size as a cue to distance: Static localization, this JOURNAL, 64, 1951, 54-67; A. H. Hastorf, The influence of suggestion on the relation between stimulus size and perceived distance, *J. Psychol.*, 29, 1950, 195-217.

[2] William Epstein, The known-size-apparent-distance hypothesis, this JOURNAL, 74, 1961, 333-346; W. C. Gogel, B. O. Hartman, and G. S. Harker, The retinal size of a familiar object as a determiner of apparent distance, *Psychol. Monogr.*, 71, 1957, (No. 442), 1-16.

[3] J. C. Baird, Retinal and assumed size cues as determinants of size and distance perception, *J. exp. Psychol.*, 66, 1963, 155-162; Gogel, The size cue to visually perceived distance, *Psychol. Bull.*, 62, 1964, 217-225.

[4] Epstein, The influence of assumed size on apparent distance, this *Journal*, 76, 1963, 257-265.

E through *O*'s hand. The length of the rope which was drawn could be read on a cm.-scale. Neither the rope nor the size-comparison series was visible to *O*.

Observers. The *O*s were 45 students in an introductory psychology course. None of them required corrective lenses.

Procedure. Each *O* was assigned to one of the three standard-coin conditions in the order of his appearance in the laboratory. The procedure was divided into two parts. (a) *Training stage:* Each *O* made distance- and size-judgments of nine size-distance combinations; the three training disks placed at three distances: 100, 135, and 170 cm. The nine combinations were presented in random order under full-cue conditions. *O* matched his visual impressions of size by selecting the raised comparison-disk which was subjectively equal to the standard. For the distance-judgments, *E* pulled the rope through *O*'s hand until *O* tightened his grip and said "stop." No effort was made to control the rate with which the rope was drawn.

TABLE I

APPARENT DISTANCE AND SIZE OF STANDARD COINS WITH MONOCULAR AND BINOCULAR VISION

(Physical distance = 135 cm.; physical size = 2.38 cm. All entries in cm.)

Standard	Apparent distance				Apparent size			
	Monocular vision		Binocular vision		Monocular vision		Binocular vision	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Dime	103.56	20.14	117.03	14.45	2.08	0.30	2.36	0.14
Quarter	128.10	27.65	123.73	12.27	2.26	0.31	2.47	0.18
Half-dollar	151.30	31.62	123.07	16.11	2.60	0.43	2.45	0.28

For both types of judgment, *O* was corrected after each trial. (b) *Testing stage:* The standard coin was presented at a distance of 135 cm. and *O* made judgments of distance and size, first under totally reduced conditions of observation, then under unrestricted conditions. The procedure for obtaining judgments was the same as in the training stage except that no corrections were made.

Results. The reason for introducing the training stage was to insure that *O*s could accurately match tactual and visual impressions of size and distance prior to the critical test. The results showed that the *O*s were able to perform at a high level of accuracy. The mean distance-judgments were 89.47 ($SD = 15.32$), 126.79 ($SD = 21.39$), and 163.21 ($SD = 25.43$) for the 100-, 135-, and 170-cm. standard distances, respectively. The mean size-judgments were 1.89 ($SD = 0.10$), 2.37 ($SD = 0.12$) and 2.69 ($SD = 0.14$) for the 1.75-, 2.38-, and 2.86-cm. standard disks, respectively.

The results for the testing stage are presented in Table I. Under conditions of monocular observation, the half-dollar (*H*) was localized most distantly, and the dime (*D*) was judged to be nearest. The perceived distance of the quarter (*Q*) was between that of *D* and *H*. Analysis of variance yielded an *F* of 11.22 ($df = 2/42$, $p < 0.001$). Duncan's mul-

tiple-range test showed that all the differences between the means are significant (Q-D, $p < 0.05$; H-Q, $p < 0.05$; H-D, $p < 0.01$). An analysis of the binocular data for distance yielded a nonsignificant F (< 1).

Table I also shows that, under the reduced conditions, H was judged to be largest, D smallest, and Q was intermediate. An analysis of variance yielded an F of 19.10 ($df = 2/42$, $p < 0.001$). Duncan's test showed that all the differences between the means were significant ($p < 0.05$). An analysis of the binocular data for size did not reveal any significant variations.

Next the size-distance relationship was assessed by means of the Spearman rank-correlation coefficient (corrected for ties). The size-distance invariance hypothesis requires a significant positive correlation between perceived size and perceived distance. The coefficients were -0.024, 0.015, and 0.66 ($p < 0.01$) for D, Q, and H, respectively.

A final analysis involved the comparison of three values for each standard: objective distance, obtained apparent distance, and theoretically required apparent distance. The latter is the value required by the known-size-apparent-distance hypothesis, and is determined by the formula: theoretical distance = (apparent size/physical size) (physical distance).[5] The predictions of apparent distance based on the known-size hypothesis are closely approximated by the data for Q and H. The mean apparent distance for D falls short of the theoretically required distance by 12%.

Discussion. The judgments of distance obtained in the present experiment conform to those of the original study. The size-judgments are also similar to those obtained previously. Since the present procedure did not involve visual comparisons, it is reasonable to conclude that the findings are not methodological artifacts of the sort described in the introductory section.

Judgments of size and distance were not always positively correlated for the *O*s under each condition. Although the expected positive correlations *were* obtained in the original study, evidence that the size-distance link was not to be interpreted too rigidly was also present in that study. Evidently, knowledge of the perceived size for a given O does not permit an exact prediction of the distance at which the O will perceive the standard. It cannot be ascertained from the data whether this failure is due to the unreliability of the measures of size and distance, or to some procedural

[5] Epstein, *op. cit.*, this JOURNAL, 76, 1963, 260.

flaw, *e.g.* availability of conflicting distance-cues, or whether these repeated failures imply an inherent limitation of the hypothesis.

Summary. Nonvisual measures of the perceived size and distance of coins were obtained. The purpose was to determine whether earlier results indicating a relationship between assumed size and perceived absolute distance were artifacts of the visual relationships introduced by the use of visual comparison-objects. The results of the present study were generally in agreement with those of the earlier study, leading to the conclusion that the assumed-size-perceived-distance relationship was not a methodological artifact.

APPARATUS

FLOATING CONSTANT-CURRENT AND CONSTANT-VOLTAGE SQUAREWAVE PSYCHOPHYSIOLOGIC STIMULATORS

By Roy J. Krusberg and Herbert Zimmer, University of Georgia

Although extensive use has been made of electric shock in studies on human pain, stress, and learning, little is known about the effects of the many stimulus-properties from which the experimenter must choose. Yet to be investigated for their relative effectiveness in producing pain are such characteristics as waveform, repetition-rates, on-off ratios, duration of rest- and recovery-periods, resistanec, capacitance, and impedance of current-pathway, etc. Even information about the primary aspects of the physical stimulus: potential, current, power, and frequency, is remarkably meager with respect to the human organism.[1] In part, the lack of accurate, inexpensive instruments may have hampered this area of research.

Constant-Current Stimulator

The stimulator was designed: (1) to avoid the introduction of ground-loops or noise which might interfere with psychophysiologic recording; (2) to be safe for use with human subjects; (3) to cover the resistance-range usually encountered with such subjects; (4) to provide acceptable regulation of current-values within this range; (5) to deliver any sequence of squarewave pulse-trains of any on-off duration the experimenter might wish to employ; (6) to respond so rapidly as not to introduce significant distortions of squarewaves supplied to it by an external source; and (7) to provide for direct reading of amperage and calculation of voltage and

* This study was conducted in the Bioelectronic Computer Laboratory, and supported by the Air Force Office of Scientific Research, Grant AF-AFOSR-257-64, and the Rome Air Development Center, Contract AF30(602)-3380.

[1] T. W. Forbes and Arthur L. Bernstein, The standardization of 60 cycle electrical shock for practical use in psychological experimentation, *J. gen. Psychol.*, 12, 1935, 436-442; B. von Haller Gilmer, The sensitivity of the fingers to alternating electrical currents, this Journal, 49, 1937, 444-449; Harris E. Hill, Harold G. Flanary, Conan H. Kornetsky, and Abraham Wikler, Relationship of electrically induced pain to the amperage and the wattage of shock stimuli, *J. Clin. Invest.*, 31, 1952, 464-472; Charles F. Dalziel, Effects of electric shock on man, *IRE Trans. med. Electronics*, PGME-5, 1956, 44-62.

wattage from subject-resistance. Although earlier stimulators[2] have been able to supply constant current, this stimulator, with hardly any additional expense, (1) provides considerably better current-regulation over a wide range of resistance-loads, (2) readily accepts input from an external pulse-generator, and (3) produces squarewave pulses up to a frequency of 200 KC.

The stimulator operates as an ungrounded system. The subject *(S)* is completely separated from the power line by an isolation-transformer. He is further protected by two 1/100-amp. fuses in series with the output jacks. Kouwenhoven *et al.* have indicated that a minimum of 50 m. amp., at 60 cycles for an on-duration of 5 sec., through the most dangerous pathway at a resistance of 650 Ω, is required to produce ventricular fibrillation in an anesthetized dog.[3] The maximal current-output of the present stimulator is limited to 5 m. amp. by the current-regulator control-circuit. The range covered by this instrument, from 0 to 400,000 Ω, is adequate for most resistance-values encountered in human subjects, whenever the electrodes are large enough and have sufficient electrical continuity with the skin to prevent burns.

Within the limits indicated in Fig. 1, current-regulation is better than 1% of the milliampere output. If one is to avoid introducing artifacts and distortions at the electrode site, it is advisable to use electrodes which are resistant to polarization, such as silver-silver chloride sponge electrodes.[4] The flexibility of this stimulator in administering any sequence of squarewave pulse-trains for any given pulse-duration or intervals between pulses is limited only by the capability of an external pulse or time-interval generator, which provides the necessary signals to it. Since the pulse-gate of the stimulator has a rise-time of less than 5 μsec., it transmits squarewave pulses with essential veracity.

Fig. 2 illustrates the simple network which is the basic current-regulation for this stimulator. The cathode of V_1 is biased to a sufficiently high voltage, by voltage divider R_2-R_4, to prevent current flow through the tube when the tap on R_3 is at ground point. As the experimenter raises the tap on R_3 from ground level until the desired current is flowing

---

[2] Robert Davidon and Nathaniel Boonin, A constant current stimulus-generator, this Journal, 69, 1956, 466-468.

[3] W. B. Kouwenhoven, G. G. Knickerbocker, R. W. Chesnut, W. R. Milner, and D. J. Sass, A-C shocks of varying parameters affecting the heart, *Communications & Electronics*, 1959 (No. 42), 163-169.

[4] Donald N. O'Connell and Bernard Tursky, Special modifications of the silver-silver chloride sponge electrode for skin recording, *Psychophysiol. Newsltr.*, 8, 1962, 31-37.

through *S*, voltage divider R_1-R_3 applies an increasing positive potential to the grid of V_1. As current increases through V_1, the voltage at the cathode of V_1 rises because of the increased IR drop through R_4. With R_3 at any given setting, the bias on V_1 will be controlled by the total cur-

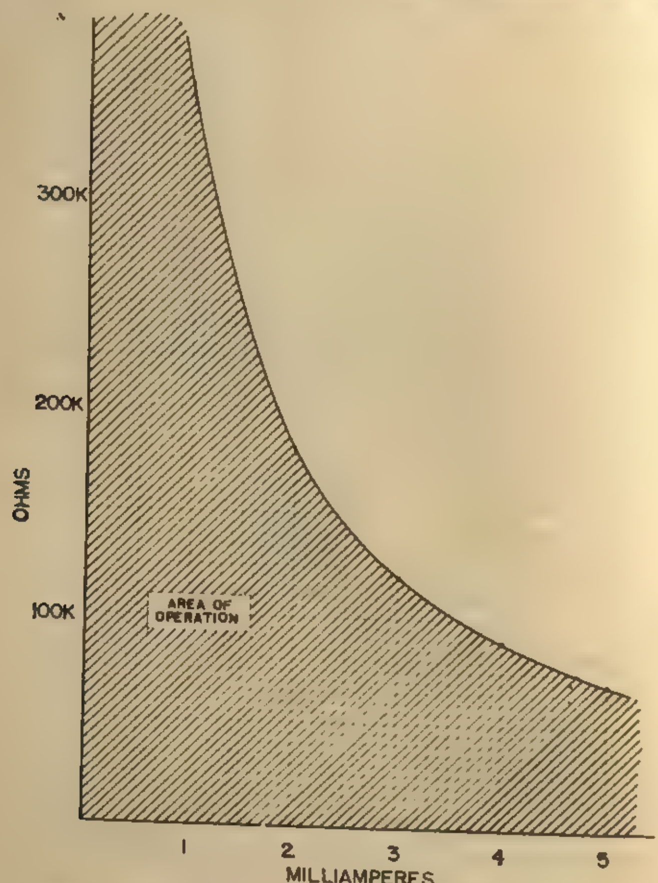

FIG. 1. CURRENT CHARACTERISTICS OF CONSTANT-CURRENT STIMULATOR UNDER VARIOUS LOADS

rent flow through R_4. Decreasing current will decrease negative grid bias, allowing more current to flow through the tube. Increasing current will increase negative grid bias, allowing less current to flow through the tube. A tube with high gain must be used for optimal regulation. The regulator-tube in this circuit is unique in that it provides both current regulation and current switching.

Additional improvements on the basic circuit include: (1) a regulated

voltage reference for current control; (2) an electronic switch for transferring current from a dummy load to the subject; (3) a meter for reading subject current; and (4) a manual current set switch. The current

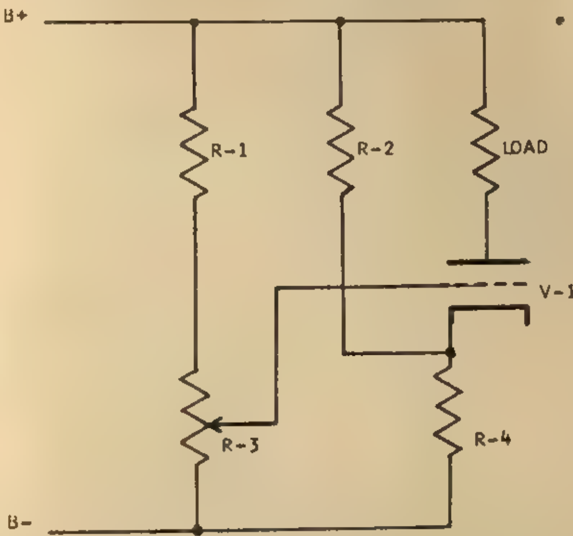

Fig. 2. Current Regulating Network of Constant-Current Stimulator

source in a 600-v. center-tapped rectifier transformer. It feeds a 6X4 full-wave rectifier, followed by a brute-force capacitor filter, in order to achieve peak transformer voltage of about 400 v.

The current-regulator and switching tube is a 6JH8 double anode sheet beam switching tube, with one anode connected to the subject electrodes and the other to a dummy load. An initial cutoff bias voltage is provided by a dropping resistor and OA2 voltage regulator tube feeding 25 ma. of current through the 1000-Ω cathode resistor. A 1N3033A zener diode across the current-control potentiometer provides a constant reference voltage for current regulation. Beam switching from dummy load with preset current to subject electrodes is accomplished by a 12AU7 dual triode bistable multivibrator, which applies a plus and minus 60 volt signal to the beam switching electrodes. Beam switching requires an external 25-v. negative pulse, or depression of the reset switch. A 26Z5W full-wave rectifier, with capacitor filter, dropping resistor and OB2 voltage regulator, supply minus 100 volts to the bistable multivibrator circuit. A separate 26.5-v. filament transformer (T-2) supplies heater current to the 26Z5W.

Constant-Voltage Stimulator

Design considerations for the constant-voltage stimulator were similar to those of the constant-current stimulators. The constant-voltage stimulator (1) is safe for use with human *Ss*; (2) covers the resistance range usually found in human *Ss*; (3) does not produce extraneous ground-loops or

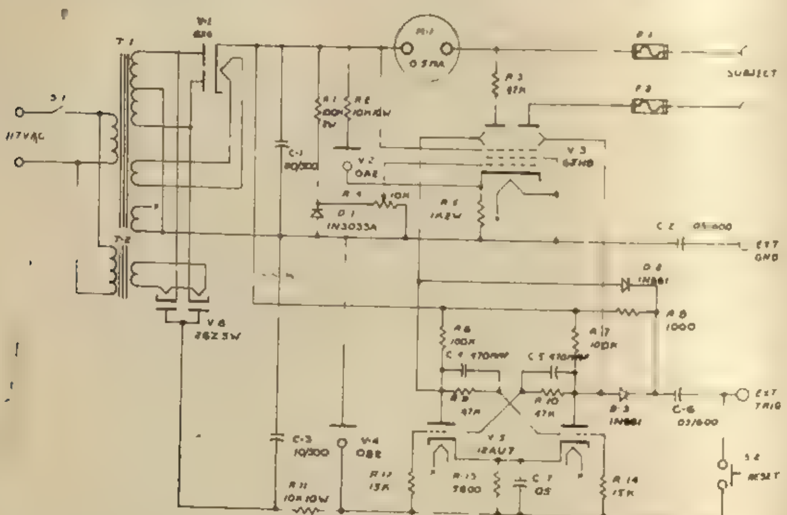

Fig. 3. Circuit Diagram of Constant-Current Stimulator

noise; (4) provides acceptable voltage regulation; and (5) permits the direct reading of voltage output, and the calculation of amperage and wattage values from measured resistance. The constant-voltage power supply detailed in Fig. 4 has a continuously variable voltage output of 10 to 150 v. Regulation is better than 1% on resistance loads down to 100,000 Ω, and better than 3% on resistance loads down to 30,000 Ω. Although the circuit illustrated in Fig. 4 shows a relay in its output stage, which limits the repetition rate to 100 c.p.s., an adaptation of a thyratron circuit to the output stage of this stimulator could produce frequencies up to 5 KC.

The voltage source is an instrument power supply isolation transformer, with a 135 v. secondary feeding a semiconductor full-wave, voltage-doubler rectifier, which provides a maximal output of 350 v. Voltage control is effected by a shunt regulator. The regulator leg is in parallel with the control leg. In the shunt regulator leg, a 6K6 power amplifier, pentode, shunt regulator tube is in series with an OA2 voltage regulator tube and a [illegible]0-Ω, 1-w. resistor. A series resistor and potentiometer are in parallel with the OA2. Also in parallel with this shunt regulator circuit is a 6AU6

voltage amplifier pentode. The 6AU6 is in series with a 56-v., 1-watt zener diode, which holds the 6AU6 cathode voltage constant, over changing tube currents. The plate of the 6AU6 is DC coupled to the 6K6 grid. The plate load resistor of the 6K6 is connected to the B+ bus. A voltage divider network supplies screen voltage to the 6AU6 and bias current to the

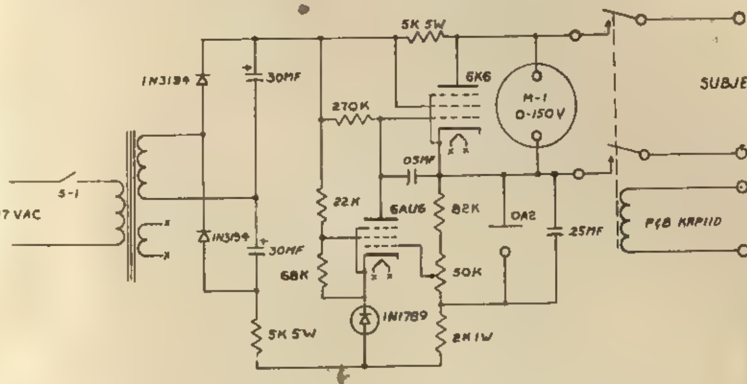

Fig. 4. Circuit Diagram of Constant-Voltage Stimulator

zener diode. The tap on the potentiometer, which is in parallel with the OA2, is connected to the grid of the 6AU6. In operation, the potentiometer tap is set to permit the desired voltage across the 6K6 shunt regulator. When a resistance load is connected across the 6K6, the increased current through the 2,000-Ω resistor raises the voltage at the resistor and at the 6AU6 grid. An increase in 6AU6 grid voltage causes an increase in current flow through the 6AU6, with consequent lowering of its plate voltage and the 6K6 grid voltage, decreasing current flow through the 6K6. Voltage across the 6K6 is thus maintained at a constant level.

Since both the constant-current and the constant-voltage stimulators were designed to be used in conjunction with psychophysiologic recording equipment, and since it is essential that *S* be grounded at a single point only, if unmanageable ground-loops are to be avoided, both units are kept entirely above ground. Some care should be exercised in the construction to isolate all components from ground, because of the potentially hazardous voltage developed by the stimulators. All circuit components are mounted on phenolic terminal boards to isolate the circuit from ground for maximal safety of *S* and minimal interference with psychophysiologic data-channels.

Total cost of parts for either the constant-current or the constant-voltage stimulator is about $30-$40 each.

A DEVICE FOR RECORDING A RAT'S POSITION IN A STRAIGHT ALLEY

By JUDSON S. BROWN and JOHN J. HOOPES, University of Oregon Medical School

In studies of locomotor conflict and of approach- and avoidance-gradients, it is desirable to obtain detailed records of the animal's position in space as a function of time. This has been accomplished by harnessing the animal to an overhead, endless-loop trolley system, connected in turn through a reduction-pulley to a linear write-out device.[1] Rough indications of spatial position have also been obtained by means of multiple photoelectric cells or electronic contact-relays located throughout the experimental space, each arranged to activate an all-or-none marker on a multiple-pen event-recorder. The first of these methods has the undesirable characteristic of offering some restraint to the animal's movements, and the resolving power of the second is poor.

An alternative method, described here, appears to obviate these difficulties. No attachments need be made to the rat, and the system will resolve movements as small as 1 or 2 in. in a total space of 72 in. The basic principle is that the force exerted by a rat on a tilting floor with a central fulcrum is directly proportional to the distance of the rat from the fulcrum. By means of a strain-gauge force-displacement transducer, the varying forces arising from changes in the rat's position on the floor can be translated into proportional displacements of an ink-writing millivoltmeter.

The essential components of the system currently being used in studies of approach-avoidance conflict are shown in Fig. 1. The floor of the alley is constructed of two parallel, *L*-shaped aluminum rails, 6 ft. long, held together by flat aluminum plates at a uniform spacing of 4 in. Narrow strips of phenolic running along the top inside surfaces of these rails serve as insulating supports for the ³⁄₃₂-in. stainless-steel bars that form the grid-floor. A flat steel plate (3-in. wide × 12-in. long × ½-in. thick) with a milled, inverted *V*-groove is attached to the center of the alley floor, which

* This work was supported by Grant MH 06900-02 from the Public Health Service.

[1] J. S. Brown, Gradients of approach and avoidance responses and their relation to level of motivation, *J. exp. Psychol.* 41, 1948, 450-465.

is separate, of course, from the alley itself. (The latter, shown only in the small-scale drawing at the lower right corner of Fig. 1, is supported on legs that straddle the grid-floor.)

The grooved plate rests on a steel knife-edge which, in turn, is solidly set into a heavy wooden base. The four-element bonded strain-gauge Grass Type FT 03) is located half-way between the fulcrum and the end of the alley. This position was determined empirically as that which yielded a full-scale deflection when a 150-gm. weight was at the extreme end of the beam. In our situation, the gauge is fitted with the 300-gm.

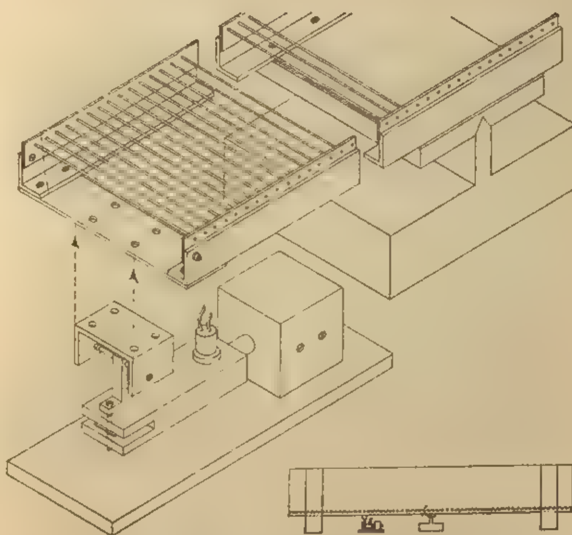

FIG. 1. THE APPARATUS

The large upper drawing (which is not to scale) shows details of the runway-floor, grid-bar supports, knife-edge floor-pivot, and strain-gauge with pivoted linkage. The small lower drawing (to scale) indicates the relation of the pivoted floor (dotted lines) to the strain-gauge and to the sides of the alley and its supporting legs.

compression springs provided by the manufacturer, and the strain-sensitive blade is connected to the floor-beams by means of a swivel-connection. This type of attachment permits movements of the floor to be transmitted to the strain-gauge without undue constraint. As the rat moves from the center of the alley to the right end of the beam, the sensitive element of the gauge is pulled upward. Movements to the opposite end depress the element. The springs within the gauge provide the only restraint to move-

ment of the beam. A 300-gm. rat produces a maximal-beam-deflection of about 1/16 in. in either direction.

The gauge, a bridge-configuration, is energized by 8 v. from a transistorized, regulated power supply (Heath Co. Model IP-20). The output of the transducer is fed into a 10,000-Ω, 10-turn helical potentiometer which serves as a voltage-divider for the 0-10 mv. Varian recorder (Model G-10). Since rats of different weights will deflect the beam by different amounts, with distance constant, the potentiometer must be adjusted to compensate for individual differences in weight. For this purpose, a calibration-curve is used which gives the required setting for any rat within

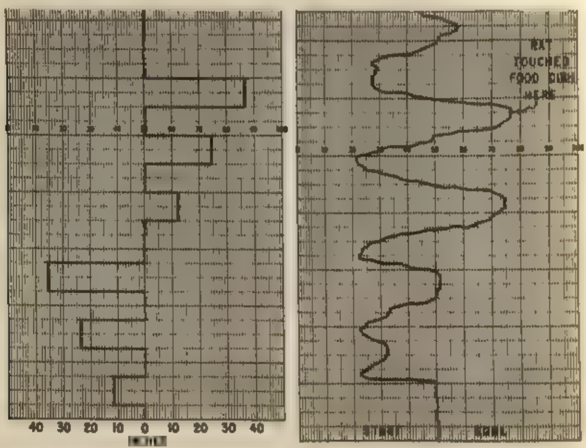

Fig. 2. Sample Tracings

The graphic record at the left indicates the response of the system to a weight of 200 gm. placed in succession 12, 24, and 36 in. to the left of the center of the alley, followed by a similar series of placements to the right of center. The record at the right shows the vacillating behavior of a rat that had been both fed and shocked at the goal-end of the alley.

the range of 150-400 gm. Thus, since the bridge-voltage remains constant, it is necessary only to weigh each rat immediately prior to testing and to set the pot to a number read from the curve. This procedure yields indications of spatial position accurate to about 2%.

The left panel of Fig. 2, reproduced from a Varian record, shows the results obtained by placing a weight of 200 gm. at distances of 12, 24, and 36 in. from the left and right ends of the balance. The curves in the right-hand panel show the vacillatory movements of a rat that has been both fed and shocked at one end of the alley. The nature of the system is such that

a leap of the rat into the air will be recorded as a movement back toward the center of the alley. Leaps usually can be distinguished from movements along the alley, however, since the former yield more peaked tracings than the latter.[2]

[2] While our system was designed for rats, the basic principle is applicable to other species, such as dog, cat, monkey, or pigeon, and to the recording of other behavioral characteristics such as general activity. Other strain gauges and other recorders could readily be substituted for those mentioned above. Investigators with limited funds might find the Heath Company's Servo Recorder (Model EUW-20) an entirely acceptable substitute for the Varian, since the price of the former is roughly one-fourth that of the latter.

APPARATUS NOTES

AN INEXPENSIVE MULTI-CHANNEL RECORDER

When more than one class of events can occur on a given trial, it is often necessary to record the sequence of events as well as the number in each class. Sequential records can be recovered from a multi-channel, pen-writing device, but most such recorders are driven by a synchronous motor that runs continuously. As a result, the intertrial spacing of events on the record is as irregular as the pace assumed by *S*, and tabulation of the record by trials is very laborious. This note describes a method by which an electric typewriter can be used to provide uniform intertrial spacing and a large number of recording channels at a relatively low cost.

In an ordinary typewriter, a system of levers is used to amplify the force applied by the typist's fingers; in an electric typewriter, a motor is used to drive the type, the same motor for all the type. Depressing a key on an electric typewriter does not close a circuit, but serves instead to engage that key in the motor, which then drives the selected type against the platen. Nevertheless, so little force and travel are required to operate the keys of an electric typewriter that they can be operated by small solenoids which can, in turn, be actuated by the same pulses that actuate cumulative counters.

We have used Guardian Electric solenoids (Type 11, 24v. D.C.) on a rack attached to the frame of the typewriter directly under the keyboard. The solenoids are attached to selected keyrods with small tension-springs which serve to protect the coils of the solenoids and the alignment of the keyrods. Mounting the solenoids from below and attaching them to the keyrods leaves the keyboard itself free for manual typing. Thus, codes and other information can be typed in without removing the data-sheet. Our machine has been in operation for approximately 15 hr. per day for more than 70 days without requiring service. (To prevent overheating, the motor should be turned off periodically for short intervals.)

Wellesley College

R. Allen Gardner
William B. Coate

A VARIABLE-FREQUENCY DRIVE FOR SYNCHRONOUS MOTORS

With an audio-oscillator and a suitable amplifier, it is possible to vary the speed of apparatus driven by synchronous motors. For example, a Gerbrands model M1-A memory-drum may be operated at speeds from about half to double its normal range of 1 to 3 presentations per sec., with a very large number of speeds in between.

The frequency-control knob of a Hewlett-Packard model 200B oscillator is replaced with a vernier-knob assembly to increase the reliability of repeated settings. The output of the oscillator may be fed directly to a 15-w amplifier with about a 500-Ω output, such as a Heathkit model A9-C. Alternatively, a conventional amplifier

* This work was supported by Grant MH 4337-02 from the National Institutes of Health. R. A. Gardner is now at the University of Nevada, and W. B. Coate at Hazelton Laboratories.

with an 8-Ω output may be modified in either of two ways: (1) by substituting a power transformed (such as Stancor PC 8401) for the usual output transformer; or (2) by using a filament transformer (such as Stancor P 3046) to step up the voltage across the 8-Ω output (about 10 v.) to the required 110 v. The second method, much less efficient than the first, requires about a 30-w. amplifier, but the connections are easier to make, since they are external to the amplifier. A 15-w. amplifier is adequate for the first method. The output terminals of an amplifier must never be disconnected while the power is on. It is a good idea to connect a 500-Ω, 2-w. resistor across 8-Ω outputs, or a 5-w. lamp across 500-Ω outputs to prevent mistakes.

RICHARD O. ROUSE
OGDEN BRANDT
Austin Riggs Center, Inc.
Williams College

A METAL GLOVE FOR PROTECTION AGAINST ANIMAL BITES

A glove used by meatcutters affords dependable protection against the bites of small animals such as rats. It is made of a metal mesh sturdy enough to prevent tooth-penetration but sufficiently flexible for ease of handling. The metal glove occasionally startles rats, but this difficulty may be remedied by wearing a light cloth glove over it. The cloth also reduces any noise the mesh might make in touching cages. In work with larger animals, such as monkeys, a mesh arm-guard may be added for protection to the arm.[1]

KARL E. THALLER
University of Connecticut

[1] A catalog of these gloves may be obtained from Whiting & Davis Company, 23 W. Bacon Street, Plainville, Massachusetts.

NOTES AND DISCUSSIONS

PARTIAL REINFORCEMENT AND THE FISH

In a recent paper, Wertheim and Singer report that 24 sessions of continuous reinforcement *plus* 24 sessions of variable-interval (*VI*) 50-sec. training produce greater resistance to extinction in the fish than do 24 sessions of continuous reinforcement *alone*.[1] Since we have already reported (in a paper cited by Wertheim and Singer) that the fish's resistance to extinction is increased by *VI* 1-min. training as compared with continuous reinforcement even when the total number of training sessions and the total number of reinforcements are the same,[2] their results come as no surprise. We are surprised only by the fact that Wertheim and Singer considered them to be worth publishing, and by the amount of misunderstanding which their paper reveals.

In the first (discrete-trials) experiments on partial reinforcement in the fish, with number of *trials* rather than number of *reinforcements* equated (it is impossible to equate both at once), resistance to extinction was found to be greater after continuous than after partial reinforcement.[3] In subsequent experiments (of both the discrete-trials and free-operant varieties), with number of *reinforcements* equated, the conventional partial-reinforcement effect (*PRE*) appeared; that is, resistance to extinction was greater after partial reinforcement.[4] How were these results to be accounted for? The interpretation we proposed was that the Gellermann schedule of reinforcement used in the early equated-trials experiments (in free-operant terms, the Gellermann schedule is *VR*-2 with no run of unreinforced responses greater than three) did not increase the resistance of the partial animals enough to offset the difference in number of reinforcements which

[1] G. A. Wertheim and R. D. Singer, Resistance to extinction in the goldfish following schedules of continuous and variable interval reinforcement, *J. exp. Anal. Behav.*, 7, 1964, 357-360.

[2] R. C. Gonzalez, R. M. Eskin, and M. E. Bitterman, Extinction in the fish after partial and consistent reinforcement with number of reinforcements equated, *J. comp. physiol. Psychol.*, 55, 1962, 381-386.

[3] Jerome Wodinsky and M. E. Bitterman, Partial reinforcement in the fish, this Journal, 72, 1959, 184-199; Wodinsky and Bitterman, Resistance to extinction in the fish after extensive training with partial reinforcement, *ibid.*, 73, 1960, 429-434; Nicholas Longo and M. E. Bitterman, The effect of partial reinforcement with spaced practice on resistance to extinction in the fish, *J. comp. physiol. Psychol.*, 53, 1960, 169-172.

[4] Gonzalez, Eskin, and Bitterman, *op. cit.*, 381-386.

favored the continuous animals. Even with equated reinforcements, we found the *PRE* to be rather small in Gellermann animals—much smaller, at least, then in animals trained with longer runs of unreinforced responses. That it should be possible in the fish as it is in the rat to produce the *PRE* without equated reinforcements (that is, with equated trials), we saw little reason to doubt; it should be necessary only to use a schedule of reinforcement which provided for longer runs of unreinforced responses.[5]

Wertheim and Singer contradict this interpretation. They suggest instead that the early (equated-trials) results were due to the fact that resistance to extinction was measured only in single post-reinforcement sessions, and they cite Keller and Schoenfeld in support of the proposition that the same results would be found in parallel experiments with the rat.[6] In their own experiment, they report, resistance to extinction was greater after continuous reinforcement in the first of a series of post-reinforcement extinction-sessions, but greater after partial reinforcement in subsequent sessions, even though they did not equate number of reinforcements or use long runs of unreinforced responses in training.

As it happens, however, one of the early experiments with the fish which showed greater resistance to extinction after continuous reinforcement *did* involve a long series of extinction-sessions—it was a discrete-trials experiment with a 24-hr. intertrial interval.[7] Both the rat[8] and the pigeon[9] show the *PRE* under such conditions. Furthermore, there is no evidence whatsoever that the *PRE* is unlikely to appear (or that the opposite effect is likely to appear) in a first post-reinforcement extinction-session if the session lasts long enough to produce a substantial amount of extinction. With number of reinforcements equated, the fish certainly shows the *PRE* in a single session, both under free-operant and under discrete-trials conditions.[10] A single session also has been adequate to show the *PRE* in free-operant work with rat[11] and pigeon.[12] Keller and Schoenfeld

[5] Gonzalez, Eskin, and Bitterman, Further experiments on partial reinforcement in the fish, this JOURNAL, 76, 1963, 366-375.

[6] F. S. Keller and W. N. Schoenfeld, *Principles of Psychology*, 1950, 89.

[7] Longo and Bitterman, *op. cit.*, 169-172.

[8] Solomon Weinstock, Resistance to extinction of a running response following partial reinforcement under widely spaced trials, *J. comp. physiol. Psychol.*, 47, 1954, 318-322.

[9] D. H. Bullock, W. A. Roberts, and M. E. Bitterman, Resistance to extinction in the pigeon after partially reinforced instrumental training under discrete-trials conditions, this JOURNAL, 76, 1963, 353-365.

[10] Gonzalez, Eskin, and Bitterman, *op. cit.*, *J. comp. physiol. Psychol.*, 55, 1962, 381-386.

[11] W. O. Jenkins and M. K. Rigby, Partial (periodic) versus continuous reinforcement in resistance to extinction, *J. comp. physiol. Psychol.*, 43, 1950, 30-40.

[12] W. O. Jenkins, Howard McFann, and F. L. Clayton, A methodological study of

say nothing about the number of sessions required to show the *PRE*, but only that the opposite effect may appear for "a brief period" after the cessation of reinforcement.[13] The single extinction-sessions of the early fish experiments lasted considerably longer than "a brief period." They were continued to the point of five successive failures to respond in 30 sec., although the latency of response at the end of training was consistently less than 1 sec. By contrast, the sessions of Wertheim and Singer were brief indeed—their first post-reinforcement sessions characteristically produced no sign of extinction—but, even so, the claim that there was a significantly greater level of response after continuous than after partial reinforcement in their first post-reinforcement sessions is without foundation. For proper balancing of order, only the last two extinctions for each animal may be compared, and in those extinctions the first-session performances under the two conditions were practically identical for three of the four *S*s (S-1, S-3, and S-4). It is encouraging to find the Skinnerians beginning to make use of statistical techniques, but now they must learn something about the conditions under which the various techniques are applicable.

Did Wertheim and Singer produce their results without using long runs of unreinforced responses in training? Although no details are provided, their sample tracings suggest that runs of at least seven or eight unreinforced responses must have occurred on occasion. That there were runs longer than three (the Gellermann maximum) cannot be doubted. But Wertheim and Singer are right about one thing: they did not equate number of reinforcements. Nor, for that matter, did they equate number of training sessions. What they compared was resistance to extinction after 24 sessions of continuous reinforcement *plus* 24 sessions of *VI*-training with resistance to extinction after 24 sessions of continuous reinforcement alone, and one cannot say, therefore, whether the difference in over-all resistance which appears in their experiment is due to the difference in schedule, to the difference in number of reinforcements, to the difference in number of training sessions, or to some combination of these. Suppose that partial reinfocement had not been studied prior to the experiment of Wertheim and Singer. Would we have been willing to conclude from their results that VI produces greater resistance to extinction than does *CRF?* Certainly not. In our opinion, their results have no meaning at all.

The suggestion in our 1963 paper that the mechanisms responsible for the *PRE* in fish and rat may not be identical was based, not on the results

---

extinction following aperiodic and continuous reinforcement, *J. comp. physiol. Psychol.*, 43, 1950, 155-167.

[13] Keller and Schoenfeld, *op. cit.*, 89.

considered by Wertheim and Singer, but on the results of experiments with alternating reinforcement, with partial delay of reinforcement, and with partial reinforcement in widely spaced trials, which Wertheim and Singer entirely ignore. All of the results thus far obtained with the fish are compatible with the sensory-carryover interpretation of Sheffield,[14] which clearly is inadequate at the level of the rat. Further evidence on this question will be published in this JOURNAL.[15]

Bryn Mawr College

R. C. GONZALEZ
M. E. BITTERMAN

ACCOMMODATION AS A CUE TO DISTANCE

Wallach and Norris recently reported an experiment that led them to conclude that "accommodation can function as a potent cue to distance."[1] We wish to point out some difficulties with their interpretation of this experiment and of its relation to the experiments of Heinemann, Tulving, and Nachmias.[2]

In Wallach and Norris' experiment, *O* equated the size of two squares. One of these (the variable) was presented at a distance of 38 cm. from *O*'s eye, while the other (the standard) was presented at a distance of 66 cm. The variable and standard squares were located near the bottom and top of one face of an oblique projection of a cube which was placed on a table-top. The scene was viewed monocularly from the perspective-point of the cube-projection. The standard and variable were presented simultaneously, and *O* made size-matches under three conditions: (1) while viewing the stimulus-forms through a peephole that was larger than the size of the natural pupil; (2) while viewing them through an artificial pupil 0.95 mm. in diameter; and (3) with the cube-projection covered, and (presumably) while viewing through the large peephole. The mean of the size-matches obtained was 3.26 cm. *with* the artificial pupil, 3.68 *without* the artificial pupil, and 4.17 cm. with the cube-projection covered. The authors conclude that the difference obtained between the first two conditions reflects the effect of accommodation on apparent *distance*. The following are some of the objections that may be raised to this conclusion.

[14] V. F. Sheffield, Extinction as a function of partial reinforcement and distribution of practice, *J. exp. Psychol.*, 39, 1949, 511-526.

[15] R. C. Gonzalez, E. R. Behrend, and M. E. Bitterman, Partial reinforcement in the fish: Experiments with spaced trials and partial delay, this JOURNAL, 78, 1965 (in press).

[1] Hans Wallach and C. M. Norris, Accommodation as a distance-cue, this JOURNAL, 76, 1963, 659-664.

[2] E. G. Heinemann, Endel Tulving, and Jacob Nachmias, The effect of oculomotor adjustments on apparent size, this JOURNAL, 72, 1959, 32-45.

(1) The assumption that accommodation is the variable that governs the change in the size-matches is not justified. Even in monocular vision, changes in accommodation are accompanied by changes in convergence. Whether changes in accommodation alone can produce changes in apparent size is not yet known, but the experiments of Heinemann, Tulving, and Nachmias showed that changes in convergence, unaccompanied by changes in accommodation, are sufficient to produce changes in apparent size of the magnitude found by Wallach and Norris.

(2) Wallach and Norris measured changes in size-matches but their conclusions refer to *apparent distance.* In inferring changes in apparent distance from their results, they make two assumptions. The first is that the difference in the matches obtained under the first two conditions reflects a difference in the apparent size of the standard. This assumption will be discussed in Section (3) below. The second is the time-honored assumption that apparent size is proportional to apparent distance. The validity of this second assumption is questionable.[3] Under monocular viewing conditions, Bappert and, more recently, Heinemann, Tulving, and Nachmias found a relationship between apparent size and distance directly opposite to that assumed in the traditional size-distance hypothesis.[4] Since the conditions of Wallach and Norris' experiments differed in some ways from those of Bappert and of Heinemann, Tulving, and Nachmias, it is possible that in their experiments apparent distance did vary with apparent size in the manner they postulate, but that can hardly be taken for granted.

(3) The results do not necessarily indicate that the standard square has a larger apparent size when viewed with the natural pupil than when viewed with the artificial pupil. The trouble is that the comparison-(variable) square was not viewed under constant conditions. When the standard was viewed with the natural pupil, it was matched to a variable viewed with the natural pupil, and, when the standard was viewed through the artificial pupil, so also was the variable. The change in the size of the pupil almost certainly affects the apparent size of the variable as well as the apparent size of the standard. All that the experiment shows is that the *match* changed as a function of pupil-size. In other words, a change in the size of the pupil affected the apparent sizes of the variable and standard differentially. Whether either the standard or variable remained constant in size while the

---

[3] William Epstein, John Park, and Albert Casey, The current status of the size-distance hypothesis, *Psychol. Bull.*, 58, 1961, 491-514.

[4] Jacob Bappert, Neue Untersuchungen zum Problem des Verhältnisses von Akkommodation und Konvergenz zur Wahrnehmung der Tiefe, *Z. Psychol.*, 90, 1922, 167-203; Heinemann, Tulving, and Nachmias, *op. cit.*, 40 ff.

size of the other changed, or whether both appeared larger or smaller under one condition, is not known.

(4) Referring to the experiments of Heinemann, Tulving, and Nachmias, Wallach and Norris state:

> These authors obtained size-matches with accommodation the sole cue for distance and found not more than 7% size-change for a distance-change equivalent to one diopter of accommodation. In our experiment, a distance-change equivalent to 1.1 diopters caused a size-change of 13%, and this effect is achieved in the process of counteracting perspective cues to distance. The discrepancy is, however, readily explained. The earlier authors [Heinemann, Tulving, and Nachmias] created conditions unfavorable to demonstrating an effect of accommodation as a cue for distance by placing the variable object at a distance of 4 m. from *O*. At such a large distance, accommodation cannot be expected to operate effectively.[5]

In the experiment of Heinemann, Tulving, and Nachmias, a number of test-objects placed at various distances well within the range within which accommodation is highly accurate were matched to the variable. While it might be argued that the position of the variable at a distance of 4 m. from *O* could make for high variability in the size-matches, it is difficult to see how the distance of the variable could affect the measured rate at which size changes with distance. Recently, Biersdorf, Ohwaki, and Kozil have reported similar size-changes with monocular viewing and 'analytic' instructions.[6]

Actually, the comparison made by Wallach and Norris is invalid. The change of 13% found by these authors refers to the difference in the setting of the variable under the following two conditions: (a) both variable and standard viewed with the natural pupil; and (b) both variable and standard viewed with the artificial pupil. As already noted, this difference cannot be interpreted in any straightforward way as a difference in the apparent size of the standard. On the other hand, the change of 7% per diopter of accommodation found by Heinemann, Tulving, and Nachmias was obtained under conditions in which all observations were made with the natural pupil. Since these same authors also made similar measurements using an artificial pupil of 0.5 mm. in diameter, it is possible from their data to compute the difference that is comparable to that computed by Wallach and Norris. For a test-object placed at the distance of the standard in Wallach and Norris' experiment, the difference in the matches obtained with and without the artificial pupil is about 11%. The remain-

[5] Wallach and Norris, *op. cit.*, 663 f.

[6] W. R. Biersdorf, Sonoko Ohwaki, and D. J. Kozil, The effect of instructions and oculomotor adjustments on apparent size, this JOURNAL, 76, 1963, 1-17.

ing discrepancy of 2% is surprisingly small in view of the large individual differences found in such experiments and of the differences in experimental conditions: Wallach and Norris' *O*s saw standard and variable simultaneously and in the presence of a cube-projection, while Heinemann, Tulving, and Nachmias' *O*s saw the test-objects successively and in the absence of other visual stimuli.

Brooklyn College — Eric G. Heinemann
University of Pennsylvania — Jacob Nachmias

APPARENT DISCRETE MOVEMENT

In the course of a recent study of apparent continuous movement, the phenomenon of apparent discrete movement was discovered.[1] The work was done with standard 16-mm. film and a standard 16-mm. projector. The film was completely transparent except for an opaque line extending the full height of each frame, with this line translated on successive frames across the width of the film in equal increments. The resultant stimulus on the screen consisted of a vertical black line moving in a saltatory fashion from left to right across a white background at the projection-rate of 24 frames per sec. and with a fixed displacement-increment. The width of the line, its displacement-increment, and the background-luminance were varied systematically. The subject sat 12 ft. from the screen.

When the displacement-increment was small, the vertical line appeared to move continuously from left to right (the *phi*-phenomenon); as the displacement-increment was increased, the line appeared to move discretely. When the displacement-increment became greater than three times the line-width, it was clear that the number of apparent discrete locations was three times the number of actual discrete locations of the physical stimulus.

The number "three" is equal to the number of times the line was projected at each location on the screen. This repetition at each location is provided by standard motion-picture projectors to achieve a flicker-frequency above the normal flicker-fusion threshold. With the projector used, each frame-cycle consisted of three exposure-phases of 8.5 m.sec. each, with each exposure-phase followed by a dark phase of 5.4 m.sec. During the third dark phase, the next frame was positioned in preparation for its first exposure.

Under the testing conditions employed, apparent discrete movement was

---

[1] Edward Levonian, Perceptual threshold of discrete movement in motion pictures, *J. Soc. Motion Picture and Television Engineers*, 71, 1962, 278-281.

noted when the displacement-increment was greater than the product of the line-width and the number of repetitions of the stimulus at each physical location. The phenomenon occurred for two values of line-width and five values of background-luminance. Presumably the phenomenon would also have occurred if the experimental conditions had involved a white line on a black background, since the reversal of brightness has relatively little effect on apparent movement.[2]

In effect, then, the subject generated two of three perceived lines. Every third perceived line corresponded to a physical stimulus, but these lines appeared not to differ in quality from pairs of intermediate lines, and all lines appeared equally spaced. The *S*s were not aware that two out of three lines had no physical counterpart, and even when they were appraised of this fact and encouraged to identify the one impression of three corresponding to a physical stimulus, they were unable to make such an identification above the chance-level.

This phenomenon may have a bearing on the Gestalt explanation of apparent continuous movement in terms of the interaction of electrical fields. According to this explanation, each physical stimulus (such as the line on the screen) initiates an electrical field in that part of the cortex associated with retinal stimulation. When one vertical line is followed by a second displaced slightly in time and space, the resultant of the two cortical fields might be a continuous, unimodal distribution whose maximum undergoes a smooth spatial movement which forms the basis of apparent continuous movement. Such a mechanism does not admit field-discontinuities which the phenomenon of apparent discrete movement requires of an hypothesis based on isomorphism. Given temporal, spatial, and intensive conditions which are such as to preclude apparent continuous movement, the Gestalt hypothesis would predict either simultaneous or successive discrete impressions of stimuli only at locations associated with the physical stimuli; the hypothesis would not predict the intermediate discrete impressions obtained.

What seems to be required is a process which transforms temporally-discrete events into spatially-discrete impressions. It is conceivable that such a transformation could take place in the brain, for similar transformations do occur in nature. For instance, when a time-varying phenomenon is repeatedly sampled in such a way as to be represented as a sequence of temporally-discrete events, the sequence is represented in the frequency-domain by a spectrum which is replicated at equidistant intervals. In a man-

[2] H. R. De Silva, An experimental investigation of the determinants of apparent visual movement, this JOURNAL, 37, 1926, 469-501.

ner analogous to such a Fourier transformation, the brain may transform a restricted class of temporally-discrete events into spatially-discrete impressions.

University of California, Los Angeles EDWARD LEVONIAN

INVARIANCE OF SIZE-DISTANCE ESTIMATES

The invariance of size- and distance-estimates is demonstrated when the ratio of the two estimates (size-distance) is the same as the ratio of the size of the physical object to the distance at which it is viewed (retinal angle expressed as a tangent-function).[1] If *E* varies the retinal angle from trial to trial, the concept of invariance is satisfied if the ratio of estimated size to estimated distance varies correspondingly. Presumably, perfect invariance is demonstrated when the regression of the observed ratios (estimated size vs. estimated distance) on the physical retinal angles produces a coefficient of 1.0 and the y-intercept is zero. A regression-coefficient of 1.0 with a y-intercept $\gtrless 0$ would imply a constant error in registering the retinal angle but would, nevertheless, support the theory of invariance.

TABLE I

RATIOS OF ESTIMATED SIZE (S) TO ESTIMATED DISTANCE (D) FOR TWO VIEWING CONDITIONS AND NINE RETINAL ANGLES

Physical S/D	Retinal angle	Viewing conditions	
		reduced	unrestricted
2/360	0.0056	0.0083	0.0049
2/240	0.0083	0.0109	0.0070
3/360	0.0083	0.0106	0.0078
4/360	0.0111	0.0140	0.0106
3/240	0.0125	0.0151	0.0113
4/240	0.0167	0.0178	0.0142
2/120	0.0167	0.0170	0.0159
3/120	0.0250	0.0266	0.0221
4/120	0.0333	0.0350	0.0311
Regression-coefficient*		0.948	0.926
y-intercept*		0.014	−0.00026

* All calculations were based on six decimal places.

The theory of invariance does not in itself require a demonstration of constancy. For a given retinal angle, there always will be a single combination of size- and distance-estimates that will be correct, but there is no requirement in the theory that *O* must report that combination.

[1] William Epstein, The known-size-apparent-distance hypothesis, this JOURNAL, 74, 1961, 333-346.

In a recent test of whether size- and distance-estimates conform to the concept of invariance, Over concluded on the basis of his data that they do not.[2] If one examines the mean data which he publishes, however, one may come to the opposite conclusion. Over reported mean size- and distance-estimates for nine combinations of three distance-levels (120, 240, and 360 in.) by three size-levels (2, 3, and 4 in.) under reduced and unrestricted viewing. In angular terms, expressed as tangent-functions, his stimuli were visual angles of varying magnitudes, calculated as physical size vs. physical distance. These are given in Table I. Correspondingly, his mean responses in angular terms (see Over's Table I) were: estimated physical size vs. estimated physical distance. These observed ratios (responses) are given in Table I.

The relationship between his dependent and independent variables for the two viewing conditions can be described by regression-coefficients and y-intercepts, which are also given in Table I. The relationship, with mean regression-coefficients of 0.95 for reduced and 0.93 for unrestricted viewing, and with y-intercepts of 0.014 and -0.00027 for reduced and unrestricted viewing, gives almost perfect support for the concept of size-distance invariance. Over's method of reporting his individual data makes it impossible to determine whether they, too, would have yielded such results.

Dartmouth College HOWARD R. FLOCK

POST-RETINAL VISUAL STORAGE

If a simple outline picture or figure is placed behind an opaque cover containing a narrow slit, and if the picture is moved past the slit while the *slit itself remains immobile* with respect to the eye, then, under proper conditions, O will see the picture as a whole appear briefly in the vicinity of the slot, Fig. 1. Given proper illumination and contrast, together with an optimal speed of movement—my preliminary observations indicate that the total time to traverse the slot should be on the order of ¼–½ sec.—the figure will appear even though the slot is extremely narrow relative to the width of the figure. For example, I have successfully employed slots as small as 1/64 in. with figures more than 1 in. in width.

The significance of this effect for the understanding of short-term storage is considerable.[1] Its main point is that the whole figure is seen even

[2] Ray Over, Size and distance-estimates of a single stimulus under different viewing conditions, this JOURNAL, 76, 1963, 452-457.

[1] George Sperling, The information available in brief visual presentations, *Psychol. Monogr.*, 74, 1960 (No. 498), 1-29.

though, in fact, only a small part of it is present at any single instant. The earlier parts of the picture must then be temporarily stored, and a question of interest is where this storage occurs.

The temptation is to conclude immediately that the storage cannot be in the retinal receptors, since the slot is immobile with respect to the eye and the various 'slices' of the figure therefore are superimposed on one narrow strip of the retina. In fact, however, it is possible that the eye is 'tracking'

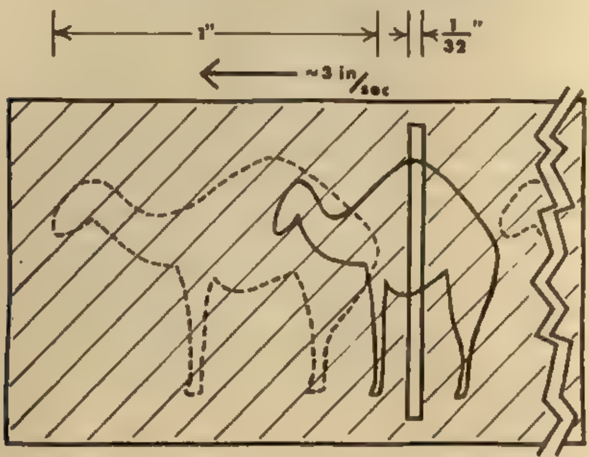

FIG. 1. PASSING A CAMEL THROUGH THE EYE OF A NEEDLE
A simple outline figure (dashed line) is passed behind a slit in an opaque screen. If it passes in approximately ½ sec., it will be seen briefly in the region of the slit, to move slightly, and to be foreshortened (solid line). The foreshortening increases with the speed of its passage until only a blur is seen. At slow speeds, the figure may be identified but will not be seen as a whole.

the figure in such a way as to 'paint' the figure in a regular fashion across the retina. Evidence against 'painting' and therefore against the notion that the storage is in the receptors—comes from the fact that the same effect occurs if two figures are moved simultaneously and in opposite directions past a single slit. Obviously, the eye cannot be tracking both figures at once.

Since the storage involved in this effect is in the form of a positive after-image, and since it seems to decay in somewhat less than ½ sec. it may well be identical with the "visual storage" discussed by Sperling[3]

[3] Sperling, *op. cit.*, A model for visual memory tasks, *Hum. Factors.*, *5*, 1963, 19-31.

and by Mackworth[3] in their research on short-term memory. If so, that component of storage is more central than the receptors and it possesses dynamic properties (the various slices of the figure must be reassembled in terms of some sort of time-of-arrival coding). It may, therefore, be of greater interest than previously assumed. Furthermore, the fact that at lower speeds *O* may report the nature of the stimulus without reporting that he has seen it directly as a whole may provide the basis for an operational distinction between visual storage and short-term storage proper.

University of California at Davis THEODORE E. PARKS

RECOVERY OF SIGHT IN THE BLIND

The psychologist confronted with the cures of sensory-deficit reported in the Gospels generally invokes an explanation in terms of hysterical conversion-symptoms removed by prestige-suggestion. The cure of the blind man of Bethsaida is of some interest because of the report furnished by the patient himself.

> They arrived at Bethsaida. Then the people brought a blind man to Jesus and begged him to touch him. He took the blind man by the hand and led him away out of the village. Then he spat on his eyes, laid his hands upon him, and asked whether he could see anything. The man's sight began to come back, and he said, "I see men; they look like trees, but they are walking about." Jesus laid his hands on his eyes again; he looked hard, and now he was cured so that he saw everything clearly.[1]

Christ's initial attempt appears to have brought about a partial cure: the man could now see but was apparently unable to distinguish between men and trees except in terms of movement. Compare the following account taken from the Getaz case reported by Von Senden.

> Joan had felt of the trunks of trees and of the trunks of her parents and supposed that they looked very much alike. They had a round trunk like a man and limbs that stuck out like arms and ended in leaves instead of hands, but if she had remained blind she would have gone through life with the vague impression that the tallest tree was about ten feet high. One of the important pieces of information that she imparted to a blind friend was this discovery that men do not really look like trees at all.[2]

Von Senden comments that trees and men belong, for the blind, to a large group of objects (*e.g.* candelabra, rounded cloak stands) which have in common the same structural schema. The confusion experienced by the

[3] J. F. Mackworth, The visual image and the memory trace, *Canad. J. Psychol.*, *16*, 1962, 55-59; The relation between the visual image and post-perceptual immediate memory, *J. verb. Learn. verb. Behav.*, *2*, 1963, 75-85.

[1] The Gospel according to Mark, 8, 22. *In the New English Bible: New Testament.* O.U.P., 1961.

[2] Morris von Senden, *Space and Sight*, Translated by Peter Heath, 1960.

blind man of Bethsaida upon first recovering his sight seems to suggest that there was a carry-over of a tactual schema to the sighted condition. It certainly implies that the man had been blind for a considerable time. I have been unable to trace any comparable report in cases of long-lasting functional disorder.

Trinity College, Dublin. DEREK W. FORREST

ERRATUM

Line 19, page 637, of the December, 1964 number of this JOURNAL, should read:

Bugelski's hypothesis was tested in the present experiment by comparing. . . .

The error did not occur in the galley proof. It was made during pagination on an otherwise errorless page.

K. M. D.

Harold Schlosberg: 1904-1964

Harold Schlosberg, Edgar F. Marston Professor and Chairman of the Department of Psychology, Brown University, died on August 5, 1964 in his 61st year, after a brief illness. In his death, psychology lost one of its sage and beloved counsellors and Brown University one of its outstanding and long-time faculty members, whose colleagues and students were his friends.

Schlosberg was born in Brooklyn, New York, January 3, 1904, the son of Ira and Harriet Schlosberg. His father was an electrical contractor and in his youth, Harold served a part-time apprenticeship as a helper in installing electrical circuits. This practical experience he put to good use in his later years in the design of, as well as general know-how with, laboratory apparatus. He attended high school in Brooklyn and entered Princeton in 1921, where he was an outstanding student. He was active in the R.O.T.C. Field Artillery Unit which in those days was still horse-drawn, and he was a member of the unit's polo team. In later years, his lectures in physiological psychology were sometimes enlivened by descriptions of the equestrian's practical use of head and neck reflexes in controlling the horse's body and stepping movements.

At Princeton, Schlosberg received the A.B. in 1925 with highest honors in Psychology, the M.A. in 1926, and Ph.D. in 1928. His was probably one of Princeton's most outstanding graduate classes in psychology for it

included as fellow students Charles W. Bray and James J. Gibson. Schlosberg's Ph.D. thesis on the conditioned patellar reflex in man is classic. It was one of the earlier systematic studies of conditioning in man in this country. A. J. Twitmyer made a chance observation of the phenomenon in 1902 and described it briefly after conditioning six subjects. It remained for subsequent workers, among them Harold Schlosberg, to give conditioning and the conditioned response its proper place in psychology.

From his association at Princeton with E. B. Holt, the behavioristic philosopher, came his appreciation and respect for the objective experimental method. There also he met Leonard Carmichael, then a young faculty member, who left Princeton in 1927 to become Chairman at Brown University. Schlosberg joined Carmichael in 1928 to initiate the modern era in psychology at Brown and from then on he was identified with psychology at Brown throughout the department's growth, first under Leonard Carmichael, then under Walter S. Hunter, and finally under his own wise and benevolent guidance.

While at Princeton, Harold met Elizabeth Fielder. They were married in 1929, a year after he joined the Brown Faculty. Their daughter Janet was born in 1934.

In the early years at Brown, Schlosberg continued active research on conditioning, on the further analysis of the factors related to the ease of conditioning in man, and on the character of conditioning in the white rat. The Department was then housed in two wooden frame buildings on the edge of the campus. His experiments on conditioning animals were set up in the attic under the eaves in one of these buildings. With this project, his genius for gadgeteering came into its own. He devised a harness to hold the rat and fabricated devices to record stimulus-onset and offset, leg movements, and breathing responses. Since this work antedated the age of transducers and remote polygraph recording, his system, making apt use of tubes, levers, string, and rubber bands, was indeed ingenious. His analysis of the differences in conditioning in man and animal and particularly his insightful paper on the relationship between success and the laws of conditioning in the *Psychological Review* in 1937 climaxed this first period of a scientific career that was to range over a wide front. He was elected to the Society of Experimental Psychologists in 1939.

During these years, he not only continued with active teaching but also did yeoman service on faculty and university committees. He served as a member of the original Board of Counselors at Brown for the undergraduate College and he was a member of the Graduate Council during the pe-

riod when the graduate program in psychology was developing and gaining strength. The first Brown Ph.D. in Psychology was awarded in 1933.

The new colleagues from Clark University who came to Brown when Walter Hunter succeeded Carmichael as Chairman in 1936 looked to Schlosberg for counsel. At that time his teaching was focused on the undergraduate course in experimental psychology and graduate seminars in learning and perception. He developed research and publication on reaction-time as a measure of response-strength. In J. McV. Hunt, as faculty colleague, he found an opportunity for collaborative research on general activity.

The second World War interrupted academic tranquility and Schlosberg, as Acting Chairman, aided by J. McV. Hunt and Carl Duncan, weathered the acceleration of the three-semester college year and the many other demands upon the small faculty group. Nevertheless, he found time to help the local Draft Board in Providence with its more difficult cases of personnel selection. Over the years he served as president of the Brown chapter of Sigma Xi and of the American Association of University Professors.

When the war ended, the veterans, many in number and high in academic motivation, returned to the undergraduate and graduate schools. A new undergraduate curriculum was developed which included psychology among the sciences and Schlosberg was largely responsible for developing the laboratory for the first course. Later (1954), the university-wide *Identification and Criticism of Ideas*, a novel seminar plan for freshmen and sophomores, was developed as a new departure in which textbooks were minimized and original sources emphasized. Again Schlosberg showed his enthusiasm for new approaches to teaching. He and the present writer developed a first course with laboratory and field trips, around the theme, "Psychology as the Science of Behavior."

Schlosberg's research continued, now in a new direction on scaling of facial expression, in two and later in three psychophysical dimensions. At this time, R. S. Woodworth was anxious to revise his *Experimental Psychology*, but the field had grown so large that he was reluctant to attempt it single-handed. Schlosberg, with his general knowledge and scholarly appreciation of a wide range of experimental psychology, was the obvious man to join him in this extensive task. The revision appeared in 1954 and immediately earned the appellation "the bible." It still remains a major-source book for experimental psychology, and just prior to his death he had begun to discuss plans for a new revision.

In 1952, Schlosberg was elected President of Division 3 of the American Psychological Association and Vice President and Chairman of Section

I (Psychology) of the AAAS. He became a Fellow of the American Academy of Arts and Sciences. He was a long time and valued member of the Eastern Psychological Association and its President in 1953-1954. He served on the editorial boards of the *Annual Reviews of Psychology* and the *Psychological Review*.

In 1954, he succeeded Walter S. Hunter as Chairman of the Department. It was just at this time that John D. Rockefeller, Jr., gave a substantial gift to Brown upon the retirement of President Henry Wriston. An allocation from these funds plus a facilities grant from the Public Health Services made possible a new building for psychology. Hunter had died shortly after relinquishing the chairmanship, and Schlosberg now became the moving spirit in the planning and supervision of construction of the new building, which was completed in 1958. The Department moved in to the Walter S. Hunter Laboratory of Psychology, at first sharing some of its then commodious space with another university research group, but a year later taking over the whole building under the increasing pressure of an expanding department, due in part to the growing strength in the traditional fields of psychophysiology and the experimental psychology of learning as well as the addition of a new program in experimental child psychology.

This period was attended by travail of Elizabeth Schlosberg's partial paralysis resulting from a cerebral shock. Schlosberg nursed her with affectionate care but ultimately she succumbed. He felt her loss deeply and there were several bleak and lonely years for him until his remarriage to the widowed Eleanor Crane Tower. He now became the new paterfamilias for the three Tower children, whom he adopted.

His concern for the affairs of the department and the university and his active involvement in them continued unabated along with an increased participation in psychological activities at the national level. He served on the APA Policy and Planning Board and chaired the APA Conference of Departmental Chairmen. In the summer of 1960, he and two colleages toured laboratories of psychology and sensory physiology in Russia and Poland for the APA on a grant from the Human Ecology Fund. He was chairman of the Panel of Consultants in Psychophysiology to the Surgeon General, U. S. Army, and had officiated at its meeting in July 1964.

His most recent research interest involved visual perception, especially the phenomenon of perceptual blanking or backward masking. With a younger colleague, he directed a series of studies which provided the data for his last scientific communication, the major invited address before the Halifax meeting of the Canadian Psychological Association in July 1964.

He was seemingly in good health at that time and thoroughly enjoyed the occasion.

This objective record gives only a hint of the man and how all who knew him remember him. His outstanding character was of unselfish personal concern for people. He was less concerned with their faults, as he looked for their strengths and brought them out. As Chairman, he guided the growth of the Department, fostering a new program in experimental child psychology and most recently a program in primate behavior. He had some misgivings about this growth, fearing that the size would vitiate the friendly atmosphere and informality that prevailed among staff and students. That the department did not lose this character is a tribute to his personality and skill. He was the stout proponent of research and scholarship. Among the faculty, he was famous for his timely quips, which not infrequently would bring aimless discussion in committees or faculty meetings to a halt.

All of this he did with little regard for his own afflictions, for over the years he suffered from arthritis. He managed so well that it was easy to forget his handicap. Fortunately, the condition became stable in his later years, although it left him with stiffened back and neck. His slightly stooped posture accentuated his naturally quizzical expression and elevated gaze which fitted his gentle smile.

Brown University CARL PFAFFMANN

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Edmund Husserls System ber phänomenologischen Psychologie. By HERMAN DRÜE. Berlin, Walter de Gruyter & Co., 1963. Pp. xv, 326.

Phenomenology of Perception. By M. MERLEAU-PONTY. Translated from the French by Colin Smith. N.Y., The Humanities Press, 1962. Pp. xxi, 466. $10.00.

Phenomenology and the Human Sciences: A Contribution to a New Scientific Ideal. By STEPHAN STRASSER. Pittsburgh, Duquesne University Press, 1963. Pp. xiii, 339. $6.00.

The Field of Consciousness. By ARON GURWITSCH. Pittsburgh, Duquesne University Press, 1964. Pp. xiv, 427. $7.95.

The Primary World of the Senses: A Vindication of Sensory Experience. By ERVIN STRAUS. Translated from the German by Jacob Needleman. N.Y., The Free Press of Glencoe, 1963, Pp. xvi, 428.

The past few years have witnessed a revival of interest among American psychologists in the philosophical foundations of their science. This is evident, for instance, in the welcome accorded the new and vigorous Division of Philosophical Psychology in the American Psychological Association. It has been signalled, too, by a significant increase in the publication, particularly in translation, of important books dealing with psychology in its more philosophical aspects. It is as though psychologists, having finally satisfied themselves that their science is now demonstrably a science, are no longer afraid to examine its fundamental assumptions, to note its weaknesses and limitations as well as its strengths, and to speculate about alternative ways in which psychology might conceivably deal more adequately with its subject matter.

Until recently the philosophers to whom the psychologists have been least reluctant to listen have probably been those associated more or less directly with the Vienna Circle. To the behavioristically minded psychologist of the 1930's what seemed to be the rigorous objectivity of logical positivism and of operationism offered an apparently adequate framework for psychological studies of behavior, particularly of animal behavior and the simpler forms of human behavior observable in rote-learning experiments. The more sophisticated forms of neo-behaviorism, as found for instance in information theory, are likewise consonant with a philosophical positivism. In continental European philosophy the positivistic movement has been paralleled since the beginning of the century by an equally vigorous phenomenological movement. Phenomenology has received relatively little attention in American and English philosophical circles, and until fairly recently it has been almost unknown to psychologists. Since with relatively few exceptions the main streams of psychological thinking in Europe have drawn more from phenomenological than from positivistic philosophy, it is important for an understanding of European psychology than the phenomenological approach as such be also understood.

The books listed above are all explicitly phenomenological. They might by some readers be categorized as philosophy rather than as psychology. They should be of

interest to psychologists, however, because their subject matter is the nature and activity of what psychologists of yesteryear were not embarrassed to call "the human mind." Phenomenologists believe they have something important to say about the mind that is not ordinarily said in the language of stimulus and response.

The most difficult of the books, and probably the least rewarding for the general reader, is the one by Drüe. It is, however, a scholarly achievement of great merit which throws light not only on Husserl's own psychological thinking but also on the thinking about psychology which took place during the early decades of this century in Germany. Husserl, while not himself a psychologist, exercised a tremendous influence on the German psychologists who were leading the protest against the associationism of the 19th century, particularly on the "schools" of Berlin, Göttingen, and Würzburg, and to a lesser extent on those of Leipzig and Vienna. This influence is evident, for instance, in the phenomenological approach of such experimentalists as Katz and Wertheimer. Husserl became embroiled, however, in the controversy over "psychologism," the claim that the laws of logic are reducible to the laws of psychology or even, in its extreme form, that all scientific laws can be so reduced. In opposing psychologism Husserl became erroneously labelled as an opponent of psychology. It is fortunate, and not merely for the clarification of the record, that among the papers now available in the Husserl Archives at the University of Louvain are a number in which Husserl deals directly and constructively with psychology's central problems as he saw them. From a combing of these sources and from a reëxamination of Husserl's earlier writings. Drüe has derived for the first time a balanced picture of Husserl as a systematic psychologist.

Of more general interest will be the translation of Merleau-Ponty's *Phénomenologie de la Perception,* originally published in 1945. Until his untimely death in 1961 Merleau-Ponty was a towering figure in European philosophy. Many would regard him, rather than Martin Heidegger, as Husserl's true intellectual successor, and certainly it is from Merleau-Ponty that the psychologist will gain the clearest insight into philosophical phenomenology as it is related to empirical psychology. Heidegger has moved far from phenomenology, and even farther from psychology, in the direction of an abstruse and almost esoteric ontology which is now of more interest to theologians than it is to scientists. Merleau-Ponty, by contrast, while sharing with Sartre and the other French Existentialists a deep concern about the ontological problem, has kept in constant contact with the observable phenomena of immediate experience. He was not himself an experimentalist, but he was thoroughly familiar with the experimental phenomenology of Katz and the members of the Gestalt group, from whom he draws liberally, and the pages of his book are studded with brilliant observations, particularly in his discussion of the body-percept, which should stimulate the experimentalist to further study.

For the reader whose frame of reference is that of the physical and the biological sciences, and the average psychologist presumably belongs in this category, Strasser is possibly an even better introduction to phenomenology than is Merleau-Ponty. Since phenomenology is an approach or a way of thinking rather than a body of doctrine, it can be understood most readily when one thinks his way with a phenomenologist through a complex problem. Strasser's question is the very old one: Can we have a unified science which includes the sciences of man (the *Geisteswissenschaften,* or "anthropology" in the broadest sense of the word). One answer, *e.g.*

Carnap's, is that the sciences can be unified by way of a single language whose basic terms are drawn from physics, the human sciences becoming thereby branches of "behavioristics." Another answer is found in the almost anti-scientist existentialism of Sartre. Strasser examines both approaches in clear and scholarly fashion, and comes out with a third alternative, his own interpretation of phenomenology, which impress this reviewer as enlightened and challenging. The social scientist will probably be disappointed at Strasser's failure to come to grips with any of the concrete problems of social research, but he will at least have a better appreciation of the philosopher's reasons for being worried about the present state of the human sciences.

The books by Gurwitsch and Straus are both translations of works which have long since established their reputations in the literature of European phenomenology. Each takes as his point of departure the challenge of the Cartesian *cogito* as reformulated by Husserl and then examines the implications of that challenge both for the empirical study of cognition and for the definition of psychology's place within the framework of philosophy. Of the two, Gurwitsch is the better oriented in the recent literature of experimental psychology, but each book deserves the attention of psychologists who are concerned about the philosophic foundations of their science.

Phenomenological exposition, because of its dependance both on the precise and on the metaphorical use of language for the communications of subtle distinctions, creates unusually difficult problems for the translator. On the whole the rendering of the above-mentioned books into English has been signally successful. The reader should remember, however, and especially as he works his way through Merleau-Ponty, that even a faithful translation frequently fails to convey the precise meaning intended by the writer. As someone put it, "Les traductions sont comme les femmes; si elles sont fidèles, elles ne sont pas belles; et si elles sont belles, elles ne sont pas fidèles."

Cornell University R. B. MacLeod

Deterrents and Reinforcements: The Psychology of Insufficient Reward. By Douglas H. Lawrence and Leon Festinger. Stanford, Calif.: Stanford University Press, 1962. Pp. 180. $4.75.

This stimulating report of research is surrounded by real and apparent paradoxes. To some readers, the largest paradox resides in the major aim of the book which is to use a limited interpretation of the human-based "cognitive dissonance theory" to explain some supposedly problematic behavior of rats. The authors have neatly cut down the sprawling dissonance model to fit the cognitive size of non-verbal animals like rats.

The particular rat-behavior which the authors set out to explain involves three seeming paradoxes in each of three empirical problems of animal learning and performance; two of which are well founded and one of which has recently been challenged by Mowrer. Why should a rat who received rewards only part of the time resist extinction longer than a consistently rewarded rat? Why should a rat who has received *delayed* reward perform longer under extinction conditions than an immediately rewarded one? And finally (and the existence of this phenomenon has been challenged by Mowrer), why should a rat who has worked harder during acquisition resist extinction longer?

The seeming paradoxes arise when we recall that more frequent rewards, more immediate ones, and less effortfully achieved ones make for stronger habits on

all measures of habit strength *except* resistance to extinction. But to make for real rather than seeming paradoxes, we must agree with the authors (to quote their last page) when they argue that their theory "inverts the traditional way of looking at such factors as non-reward, delay of reward, and effort, all of which have historically been viewed as inhibitors of behavior. Though not denying this aspect of their influence, we have regarded them as having, in addition, a facilitative effect" (p. 170). The paradox hinges upon what we take as "tradition" and "history." In reality, it is almost 20 years now that resistance to extinction has been viewed by many as a measure which involves emotional or added motivational factors, as well as "habit strengths."

Festinger and Lawrence have, however, an inherently interesting explanation of these three (or two) apparent paradoxes. The partially rewarded animal has had some trials where he has "expected" reward, has run to the goal box, and found nothing. Rather than engage in the profitless state of being frustrated, this rat (as compared to the *always*-rewarded chap) gets into a state of "dissonance" because his behavior and his expectation are embarrassingly out of whack with the fact of non-reward. The state of dissonance is a knowledge-filled and uncomfortable state which the rat attempts to reduce. He does this (or it occurs) as best he (or it) can: he finds ways of "satisfying" other motives ("subordinate" ones) by discovering "extra attractions" in the non-rewarded trial.

This is not quite what Aesop had in mind when his fox ate the sour grapes and called them sweet. Festinger and Lawrence's rat finds "*added* attractions" in that the color of the grapes is prettier, the shape of the leaves is nicer (after all, Tolman *did* grant rats esthetic needs) and besides, one can always have fun scratching oneself on a non-rewarded stop in the goal box. The same things happen when being delayed or (presumably) when working harder.

A nestling paradox here has to do with the way the authors place their contribution in the general context of learning theory. The added attractions that come from the rat version of dissonance reduction may be viewed as a new mechanism for developing "secondary reinforcers" which is not the route of secondary reward, for to quote our paradox writers: "Indeed, it is quite opposite to secondary reward in that *lack of rewards* leads to their [secondary reinforcers'] development" (p. 170).

The strategy of this work is interesting. A limited sub-set of learning phenomena are taken up, dealt with in the context of a limited human model, some 16 new related experiments are then reported, and by then there is little doubt that the present authors do better with the whole business than any competitor they deal with fully.

Yet one resists accepting into the corpus of behavior theory these bones of a cut-down human cognitive theory. Why? On the one hand, we must confess, this dissonance business might slowly expand back to its sprawling human size and take over the whole realm of animal learning. Lurking around some empirical corner, for example, is perhaps another postulated rat who will look for "additional terrors" when he gets *more* than his expectations and behavior warrant. On the other (and more probable) hand, there is a good deal of other behavioral data that are not yet integrated by dissonance considerations, and there is even the long (but not ancient) tradition of theory which may not only deal with the

trinity of problems to which Festinger and Lawrence do service, but also most of their new data.

Mowrer has recently (*Psychol. Rec.*, April 1963) forcefully restated his own theory in regard to what happens to the rat who is rewarded too little, delayed in his reward, or even (though here he has doubts) overworked. In brief, Mowrer's partially rewarded rat *does* go into a state of frustration which does not lead to strengthening the "habit," but rather, through adaptation, *neutralizes* the strength of the later frustration that arises under conditions of extinction. This partially frustrated rat (who is also partially rewarded) maintains hope (secondary reinforcers) when extinction occurs, whereas the 100% reward rat goes into extinction open to the loss of hope which frustration, which has *not* been adapted out, entails.

Mowrer's rat, if he is cognitive, is the stereotype of the stoic in that a bit of trouble during acquisition serves one in the really bad day. Serves one, that is, to persist in the face of worse trouble. Other factors will decide whether persistence was wise. If Mowrer's rat, in his cognitive incarnation, is a stoic, then Festinger and Lawrence have posited a rat with some of the properties of Pollyanna, the Glad Girl, who always differentiated out the sunny lining in every human cloud, the glorious color of every sour grape.

There is a second available explanatory approach which still appears to be viable, as exemplified in the "frustration-drive" theory of Amsel. Amsel suggests that the partially reinforced rat does not adapt *out* frustration, but rather, as partial reinforcement training proceeds, the internal cues of the (originally aversive) frustration-drive reaction tend to occur on trials on which the approach response is reinforced. These cues of frustration-drive thereby come to elicit, in part, approach responses during extinction. For the consistently rewarded animal, however, only the aversive effects of frustration-state cues are observed, hence quicker extinction.

R. K. Banks has recently (in press, *J. Exper. Psychol.*) shown that intermittently *punished* instrumental responses resist regular punishment conditions longer than do such responses which have never had a punishment history. This is predicted by Lawrence and Festinger, but they lack clear data. It is ironic (or is it paradoxical?) that Banks did his experiment based on Amsel's theory (which is dismissed by Festinger and Lawrence) rather than on a theory of dissonance reduction.

But this little book does us good service, though critics will desire more details on some of the experiments reported here. The book has moved Mowrer to become somewhat more explicit about his own approach, which is valuable. Further, a dissonance approach has some additional implications which are worth testing. But they have probably not done what we most need: showed the clear way for behavior theory to come to terms with cognitive phenomena.

The theoretical issue may turn out to be related to the temperament of the postulator. For ourselves, however, Pollyanna cannot do the job that calls for some tough old Greek like King Mithridates.

William W. Lambert
Cornell University

Adaptation-Level Theory. By HARRY HELSON. New York, Harper & Row, 1964 Pp. xvii, 732.

In Helson's world, psychological relativity is reduced to a single concept, adaptation-level (*AL*). An *AL* is the pooled effect of all the stimuli, each stimulus pulling *AL* toward its own value. The perceived character of the stimulus depends, in turn, upon its quantitative relationship to *AL*. Since the perceptual response varies inversely with *AL*, this is a theory of perceptual contrast.

There are *ALs* for different dimensions of experience and also for different domains within each dimension. The size of an automobile is contrasted with the sizes of other automobiles, a toy car with other toy cars. Like a Gestalt, *AL* reflects the temporal and spatial interaction between all relevant stimuli. But Helson reduces psychological relativity to a mathematical equation. He assigns each stimulus a discrete value on a psychological scale and then averages the rescaled values after first weighting each in proportion to its pulling power upon *AL*.

This general theory is clearest when applied to psychophysical experiments Helson rescales the physical values in accordance with how difficult they are to discriminate, *i.e.* logarithmically. Stimuli with differing effects upon *AL* are segregated into broad classes: series or comparison stimuli, standards, background, and stimuli from previous situations. The values within each class are weighted in direct proportion to their frequency of presentation, as in calculating a mean. The different classes are weighted empirically so that their grand mean equals the *AL* obtained from the experimental data. In a simplified example, $AL = aS + (1 - a)\overline{C}$, where S is the rescaled value of the standard, $\overline{C}$ is the mean of the rescaled values of the comparison stimuli, and a is the empirically determined weighting constant. When a standard is explicitly provided, *AL* corresponds to the traditional point of subjective equality. When there is no standard, *AL* is the rescaled value of the stimulus to which the neutral or middle category is applied. Other categories mark off equal psychological steps above and below *AL*.

Helson presents the quantitative applications of his theory in two extensive chapters "Psychophysical judgment" and "Perception." Major phenomena of both comparative and absolute judgment are described using the weighted-mean equation. The equation describes classical examples of perceptual contrast and also Hollingworth's "central tendency of judgment" which has sometimes been cited as an example of assimilation. Helson's treatment is much less confusing than the usual reliance upon assimilation and contrast as separate principles. The theory of *AL* seems most useful when employed to specify the effective perceptual stimulus, less so for characterizing the process of judgment and the trial-to-trial changes in *AL*. Helson has not yet attempted this sort of sequential analysis where the usual successive assimilation provides a difficult challenge to his theory.

Some readers will be put off by Helson's treatment of what appear to be negative instances. Thus, Garner's evidence for successive assimilation, that the same tone is judged louder when it follows a louder tone than when it follows a softer tone, is interpreted to be consistent with the theory (p. 155). Helson vacillates on whether the pooled effect of the series presented for comparison with a standard, $\overline{C}$, can be equated to the *AL* obtained using absolute judgments of the same series (cf. pp. 209 and 221). For some data, he adds an arbitrary correction factor (0.75d) which seems out of character with his theory of

stimulus pooling, as does his concession that *AL* might sometimes correspond to the median rather than the mean of the stimulus values (p. 60). And too little attention is given to even the most dramatic shifts in the empirical constants: for Woodrow's data (p. 205), the standard is weighted only 0.25; for similar data from his own laboratory (p. 214), Helson weights the standard 0.85.

Concern with these difficulties is fostered by Helson's emphasis upon quantitative evaluation. It would be unfortunate if this obscured the usefulness of the theory for describing the general trend of the data. The psychophysical and perceptual applications should be approached as illustrations of the theory rather than as quantitative tests. Although the range of experimental situations is impressive, the data from any specific situation are too sparse to provide more than a directional or qualitative test.

Comprehensive chapters on motivation, learning, thinking, social psychology, and personality illustrate the application of *AL* to the traditional problems of psychology. The bibliography contains more than 1000 references, many of them experiments whose details are presented in the text. The single chapter on learning includes discussion of such varied topics as generalization, transposition, latent learning, distinctiveness and retention, magnitude of reward, and partial reinforcement. To cite just one example: the section on generalization describes a central-tendency effect, the choice among test-stimuli being biased toward the central value of the available alternatives.

Helson treats each of the events affecting behavior as a stimulus whose pulling power upon *AL* can be empirically determined. When thinking, for example, specific problems or even ideas can become focal stimuli, analogous to the series of auditory tones presented for judgment in a psychophysical experiment. Instructions or immediate associations may be treated as background stimuli, physiological determinants as residual stimulation from previous situations. In this example, *AL* might be the level of difficulty that is judged neither too easy nor too difficult. Applying this approach to personality and abnormal psychology, Helson describes maladaptive behavior as inappropriate weighting of the different sources of stimulation.

This sort of analysis sacrifices power for generality. However, the analogy to stimulus pooling has suggested fresh insights to psychologists already familiar with specific areas, as in the Bevan-Adamson reinforcement model, the McClelland-Clark theory of affect, Heim's work on adaptation to level of test difficulty, Fehrer's on attitude scales, and the studies of conforming behavior by Blake and others.

Has the concept of *AL* been pushed too far? Perhaps. However, this book makes a strong case for pushing it further. The skeptical reader is less likely to complain that Helson's analyses are farfetched than that they are not specific enough. Any example of context effects, any instance of a response to a stimulus being affected by exposure to other stimuli, could have been included in this book without violating Helson's apparent criteria of relevance. These criteria do not require detailed identification of the stimuli for the *AL* equation. It is not, necessarily, that the identification could not be made. Rather, Helson is content to present his vision of a single explanation for all the effects of context. Considered as a quantitative theory, its evaluation must await further development and testing.

ALLEN PARDUCCI
University of California, Los Angeles

Testing: Its Place in Education Today. By HENRY CHAUNCEY and JOHN E. DOBBIN. New York, Harper and Row, 1963. Pp. xi, 223. $4.95.

This is an important book, for the authors have made a gallant and largely successful attempt to tell lay people just what the testers in our school systems are up to. As such, it is an effective antidote to the recent rash of books criticizing psychological testing for many if not all the world's ills. But if the antidote weren't needed, the book would look much less good; unfortunately, in their zeal for their own particular type of testing, the authors have failed to come to grips with many of education's more pressing present-day problems.

The authors' intended audience includes teachers, interested school board members, and parents; the book is smoothly written, with a minimum of technical detail and should be understandable to all these groups. At times, the presentation is too superficial, as in the two sentence discussion of construct validity on page 59 or the rather misleading and entirely inadequate treatment of item analysis on pages 42 ff. Generally, however, the book is scholarly and sufficiently cognizant of testing technicalities that the reader will be properly cautioned against dabbling with tests without sufficient and extensive training.

Chauncey and Dobbin trace the history of testing in the first chapter; their history is short and sweet, but it is accurate, and it should help to counteract the typical layman's view that testing has just recently popped on to the educational scene. In subsequent chapters, they interweave descriptions of how achievement and aptitude tests are developed with discussions of their use as tools in teaching and guidance and in selection and admission to college. I believe their strongest chapter is the one on using tests as tools in teaching. Here, they argue convincingly for the usefulness of tests as aids in planning instruction, evaluating the progress of learning, and diagnosing specific educational needs so as to individualize the course of elementary and secondary teaching and learning. Their chapter titled "What makes a test good?" is excellent. This short (11 pages) statement should be published as a separate monograph and distributed widely throughout our country's school systems to serve as a checklist of danger signs characterizing poor tests, and as a guide for use in test-program planning and test-purchasing.

I believe the least convincing chapter is the one on the use of tests for selection and admission. Here, the authors' seem to lose much of their objectivity, and their present position as President and Project-Director, respectively, of Educational Testing Service become more apparent. They state that "Many colleges that for years had been completely nonselective in admissions for philosophical or legal reasons—mostly colleges and universities supported out of public funds—are now employing selective procedures that more or less prevent students at the bottom of the ladder in academic ability from using up college space and facilities in an obviously hopeless try at higher education" (pages 109-110). Besides being offensive to the sensibilities of anyone concerned with individuality and individual excellence (wherever it may be), the above statement ignores a number of facts, some old and some new (many brought to light only recently in Darley's excellent monograph, *Promise and Performance: A Study of Ability and Achievement in Higher Education*), which need to be clarified. First, Darley's analysis of test-scores for entering students in 167 colleges in 1952 and in 1959 offers *no evidence* to support the widely held view

(perpetuated in the above statement) of a change in ability test-standards being applied for college admission. Secondly, the most effective predictor of college success remains today, as always, a measure of highschool achievement rather than any measure of scholastic aptitude. Thirdly, Darley's data, based on longitudinal studies of colleges in four states—Minnesota, Wisconsin, Ohio, and Texas—shows that students, both men and women, withdrawing from college differ only slightly on tests of scholastic ability from those who graduate. In contrast, the two groups differ sharply in high school achievement. It appears that successful completion of college is more a function of a student's prior preparation for college, his motivation for higher education, and his demonstrated ability to do academic work than of his status on a scholastic aptitude test. The authors' discussion of the non-comparability of different high schools and their educational programs is essentially irrelevant in view of these findings. Fourthly, the authors' statement, quoted above, and indeed their stance throughout the book implies a continuing subservience to *g*, with little heed paid to the many factors of ability identified during the last 30 years. In fact, they say on page 21 that "a significant point about all intelligence tests is that they measure only the individual's capacity for learning." Their reference apparently is to the ability to learn strictly academic subject matter rather than to the learning of such important but possibly less predictable proficiencies as motor skills, social and interpersonal effectiveness, "know-how" in the skilled trades, creativeness, artistic and musical powess, problem solving capability, etc. The most pressing educational problem today is to train vast numbers of persons to be useful and productive, who, according to the definition of Chauncey and Dobbin, probably have little or no "capacity for learning." What is needed is not only new educational methods and emphases, but also new psychodiagnostic tools to do a better job of pinpointing specialized aptitudes and abilities rather than to continue fixedly to emphasize solely the measurement of *g* or "capacity for learning."

Thus, the book would have been far more valuable had the authors chosen to discuss in much greater detail the other necessary measures of individuality—the measures of special skills and aptitudes, personality and motivational characteristics, and vocational interests. As it is, the book is primarily a discussion of the so-called academic aptitude tests making up the major elements of the BIG nation-wide testing programs. A book is needed which is less devoted to perpetuating the status quo, and which emphasizes and elucidates instead the way that psychological testing may be and is being used for assessing the individuality of each and every person in order to help in the enormously difficult task of utilizing maximally and humanely the vast reservoir of human resources in our society.

I do not mean to be unduly harsh. I suppose what I am really asking is for a book titled *Testing: Its Place in Education Tomorrow.* As a description and explanation of what is going on *Today,* the present book is good; and, as already mentioned, it *does* provide a useful antidote to the many poor books on testing that have flooded the market in recent years. Thus, you may still find this book useful for "educating" your lay friends, but when you do, be sure to supplement their education with the comment that this is just a starter and that hopefully the practices in the years ahead will be greatly improved over those described herein.

University of Minnesota MARVIN D. DUNNETTE

Restoration of Function after Brain Injury. By A. R. LURIA. Translated from the Russian by Basil Haigh and edited by O. L. Zangwill. New York, Pergamon Press, Macmillan, 1963. Pp. xiv, 277. $10.00.

Efforts directed toward restoration of functions impaired by brain injury may have more than humanitarian significance. Choice of therapy demands inquiry into the nature of the disorder, and its success or failure provides a practical test of the correctness of the hypothesis. This book, written in 1948 but only recently translated, brings to a wider audience the benefit of Professor Luria's astute analyses of disorders produced by brain injuries, together with what seem to be promising methods of inducing compensation. The material consists of cases with injuries incurred in the Second World War who were studied at the Clinic for Nervous Diseases in Moscow.

Since brain tissue, once destroyed, does not regenerate, restoration of impaired function by any means whatever requires explanation. In some disturbances, it can be assumed that the relevant tissue was not actually destroyed but merely underwent a reversible change. Professor Luria suggests that a component of this nature may enter into the symptomatic picture following any brain injury. The physician's task is to hasten the process of "de-inhibition," and he presents some amazing evidence of the therapeutic efficacy of certain drugs in pareses, sensory losses, and impaired movement sequences. If the relevant tissue is destroyed, rather than temporarily depressed, there are at least two ways in which function might be restored. The first is based on the assumption of plasticity (interchangeability of parts) of the nervous system: the function may be taken over by another structure. Professor Luria does not question this mechanism, although the experiments of Sperry and others (*e.g.* R. W. Sperry, *Quart. Rev. Biol.*, 20, 1945, 311) might have been cited as indicating that the basic notion of plasticity and spontaneous reorganization has severe limitations. Restoration of function by this means is, in any case, not a major concern of the book; its thesis concerns a second method of restoration. Professor Luria suggests that a radical reorganization of the impaired function may be brought about, which makes use of preserved neural structures not involved (he assumes) in the original form of the activity.

With respect to injuries which impair lower levels of integration, application of this principle may merely mean that the function is transferred to undamaged receptors or effectors (*e.g.* a pseudo-fovea may partially compensate for a central scotoma), although even simple sensory or motor defects may be improved by incorporation of the impaired activity into a functional system integrated at a higher level (*e.g.* simple tapping *vs.* tapping the answers to arithmetical problems). A more important application of the principle, however, is with respect to injuries causing impairment at higher levels of integration—praxis, gnosis, speech, and thinking. The first step is to identify the damaged link in the chain of functions making up the complex activity which is lost. For example, the primary disorder in motor aphasia may be apraxia of the speech-apparatus, or loss of word forms, or loss of phonematic hearing. Reëducation takes a different form depending on which of these functions is impaired. In each case, however, training is directed toward by-passing the disturbed process, usually by raising the function from a simpler and more automatic level to one which is more complex and deliberate. Only after prolonged training does the reorganized activity become relatively facile,

but even with such a limitation, the potential importance of Professor Luria's methods, and of his conception of functional reorganization, is clear.

Unfortunately, the success of these methods is documented in most instances only by accounts of illustrative cases. One misses the experimental approach to evaluation of therapy—standardization of techniques, adequate control procedures, and statistical treatment of results. One misses also the experimentalist's tentative attitude toward theories of brain function. The way the brain works is not so thoroughly understood as the tone of this book implies, and to the extent that it is not, the rational therapy that Professor Luria desires remains a goal that is as yet unachieved. His methods unquestionably deserve to be explored further, however, both from the point of view of increasing our knowledge of brain mechanisms and of alleviating the profoundly disabling conditions which may follow brain injury.

National Institute of Mental Health — JOSEPHINE SEMMES

The Psychology of Time. By PAUL FRAISSE. Translated by Jennifer Leith. New York, Harper and Row, 1963. Pp. vi, 343. $5.50.

The ways by which man adapts to temporal changes are numerous and intriguing. Employing the data in a bibliography of 570 publications on physiological, comparative, experimental, social, genetic, and abnormal psychology, and also on philosophical views, the author presents an up-to-date discussion of the psychology of time. Around the middle of the nineteenth century, a shift from philosophical and personal views to the empirical study of the accuracy with which men perceive time began to evolve.

Part I is concerned with conditioning to time which is biological and common to animals and man. Part II, the perception of time, deals with the perceived present, the threshold of perception of succession, and perceived duration. It is assumed that the higher vertebrates perceive time under much the same conditions as man. In Part III, three chapters are concerned with control over time: the temporal horizon (temporal perspective), the estimation of time, and the notion of time (man's adaptation to change).

The author gives a well-organized presentation of a tremendous number of details in his citations, which include many philosophical views as well as empirical data. The cautious reader will note that Fraisse is seldom critical of the material he cites. Almost any quoted generalization, verified or not, is accepted as a law. Examples: law of centration; the same space of time seems shorter as we grow older; duration of all tasks is estimated to be greater by women than by men; law of greatest utility.

Fraisse appears convinced that man possesses an internal clock, a physiological time-clock, a certain sense of time independent of objective cues. Many instances are cited in an attempt to prove the existence of this internal time-clock. One of these will be enough to show the bias in interpreting such data. Subjects awakened between midnight and 5:00 A.M. were asked what time it was. The average error of judgment was 50 min. This, contrary to the conclusions of the experimenters, was considered by Fraisse to be clear evidence of temporal orientation. In the experiment, the conscious cues on which the judgments were based were written down immediately after estimating the time. These cues were considered by the experimenters as sufficiently definite to make unnecessary the need of assuming an uncon-

scious mechanism of timekeeping. This conclusion seems reasonable in view of the large mean error in the estimated times. The author is equally uncritical in citing other so-called evidence of a physiological time-clock.

Much space is devoted to the 'indifference zone' or 'interval,' a duration for which there is no systematic error of estimation. The duration of this zone tends to be approximately 0.6 to 0.8 sec. Shorter intervals are generally overestimated and longer intervals underestimated. Contrary to Fraisse's statements, the concept of an indifference zone fixed at around 0.7 sec. is not generally accepted. The so-called indifference interval varies markedly with the method of estimation and from one study to another. Seldom have any two investigators found the same value. The range extends from 0.36 to 5.0 sec. with some concentration between 0.5 to 0.7 sec. Even under fixed experimental conditions, there appears to be no single indifference interval valid for all subjects. Obtained averages from sizable groups yield values of 0.6, 0.7 and 0.9 sec., but averages rather than individual results modify the meaning of the indifference interval. In view of the variation in the results obtained in adequately designed experiments, validity of the concept of an indifference zone or interval may be questioned.

This systematic treatise on the psychology of time is the most complete and up-to-date discussion of the subject in print. The translation is very well done. Although the author has certain emphases that need to be viewed critically, as a whole the book contains an excellent and comprehensive survey of the field. Also, for the interested person, the extensive bibliography is worth the cost of the book.

Santa Barbara, California MILES A. TINKER

Method in Experimental Psychology. By GEORGE H. ZIMNY. New York, The Ronald Press Company, 1961. Pp. x, 366. $5.50.

This book, explicitly written "for undergraduate students who are taking their first course in experimental psychology," covers most of the topics one would expect. If originality is impossible, the unique way in which the material is presented and emphasized nevertheless make it quite singular. It is obvious from reading his book that the author is an experienced teacher, that he is in close touch with his students, and that he is highly sensitive to their needs and potentialities. His presentation is simple and clear, stressing practical considerations and devoting unequalled attention to the mechanics of actual experimentation.

The author's arrangement and mode of presenting the subject matter, which is commendable for its originality and fresh point of view, is also the main source of criticism. To begin with, the material is presented in such a way that the student is virtually obliged to have read the whole book before he is equipped to undertake any experimenting. As, with few exceptions, the individual chapters in this book are not amenable to use out of context, this arrangement makes its use awkward in a course planned around a full semester of laboratory work. Also, the relative emphasis of different topics within the book makes it hard going for the average undergraduate. Zimny's introductory exposition of experimental method is excellent, by far the best of its kind to appear in recent years. Alas, however, it is so thorough that it occupies the first third of the book (even more, if one counts Chapter 7, "The Importance of Method," which rather unaccountably is placed in the section on techniques). One feels that a single introductory chapter could have sufficed here, particularly as it is assumed that the students using this text would

previously have been introduced to psychology as a science.

In contrast to the long and detailed general introduction, the section actually dealing with methodology is short and rather sketchy. The distinction between an experimental design and a technique of control is never explicitly drawn. The author states that "experimental method . . . requires that measurements be taken, control be exercised, and analysis be made." He devotes a chapter to each of these aspects of experimentation, but does little toward integrating them into any sort of unified concept of design. The psychophysical methods, commonly thought of as prototypical designs, are presented as techniques of measurement, with the usual controls employed in such designs presented in another chapter as separate methods. While it is certainly well for the student to know that such control factors as counterbalancing are not exclusive appurtenances of any given class of design, the apparent dichotomy between techniques of control and "methods of measuring the independent variable" is misleading to say the least. The four chapters headed "Steps in Designing and Carrying Out an Experiment" do little to clarify matters, since they are largely concerned with procedural details. The reader is therein adjured to choose and then follow "the method" with great care, but the means of planning an overall design is left to divination (or the offices of the instructor teaching the course).

The final chapters, which deal with report writing, are so extremely detailed that the reviewer was inclined to lose the train of thought. Beginning students, however, might find these minute strictures helpful.

This book contains much that is useful. Its value as a text in a standard undergraduate laboratory course may be limited, but its vivid and original highlighting of the subject matter recommend it as an excellent vehicle of review for beginning graduate students in all phases of psychology.

San Antonio State Hospital N. G. BURTON
San Antonio, Texas

Integrating Principles of Social Psychology. By JOSEPH B. COOPER and JAMES L. McGAUGH. Cambridge, Mass., Schenkman Publishing Company, 1963. Pp. xi, 320. $7.50.

The authors state that their aim is to examine some of the problems of social psychology in terms of integrating principles that focus upon the nature of the social individual, emphasize adaptive processes, and draw upon cognitive theory. Their method for effecting integration is to evaluate the content of social psychology in relation to a set of basic concepts rather than to search for internal structure. Social psychology is defined as "the behavior science that attempts to understand the individual's social behavior and experience." Definitions involving the social interaction concept are rejected. These are legitimate procedural decisions that lead to important inferences regarding the theoretical connections between social psychology and general psychology.

The authors do not list a set of integrating principles; however, they seem to rely most heavily upon the following: genetic preconditioning, imprinting, motivation, adaptive socializing experiences, and stimulus complexity. Those principles and related research findings are brought to bear upon the problems of social psychology. Among the broad classes of problems discussed are the following: continuities and discontinuities in animal societies, animal and human ecology, mo-

tivation, learning and socialization, language and communications, the individual in organized society, leadership, attitudes, and prejudice. These diverse topics are all related to a common set of integrating concepts. Relevant research is clearly described and nicely integrated into the general discussion.

The authors find little demonstrable connection between learning principles and socialization. In fact, theories of learning and perception are hardly mentioned. The omissions are significant and intentional. The authors do not state explicitly, but their omissions imply that the only connection between learning theory and the concepts used by the social psychologist exists as a pious hope rather than as an established fact. Their analysis suggests that an integration of social psychology with general experimental psychology can be accomplished only by (1) using broadly defined integrating principles, (2) rejecting formal learning theory as an integrating principle, and (3) omitting interaction processes and structures in defining social psychology. Thus, their analysis is as important for what it leaves unsaid as for what it makes explicit.

The authors may not agree with the reviewer regarding the implications to be drawn from their work. But it seems clear, by virtue of what they have omitted and rejected, that the main stream of general experimental psychology has failed to generate closely defined principles that will account for a considerable body of knowledge identified as social psychology. The strongest connections between social psychology and general psychology are found at a primitive, psycho-physiological level rather than at the psycho-social level. This connection is based primarily upon excellent experimental data rather than upon an integrated theoretical system. It would seem, then, that we must search the content of social psychology itself for any logical structure that may exist. It is not to be found by relating social psychology to any of the popular theoretical systems in basic general psychology.

This is an important book. It should be read by the general as well as by the social psychologist. It is interesting, not only for its method of analysis, but also for the conclusions to be drawn from the analysis.

The Ohio State University RALPH M. STOGDILL

Digital Computers in Research: An Introduction for Behavioral and Social Scientists. By BERT F. GREEN, JR. New York, McGraw-Hill, 1963. Pp. xii, 333. $10.75.

This book is a good introduction to the range of uses that are made of the computer by psychologists and by scientists in related fields. It begins with a description of computers and programming techniques, and ends with a discussion of the logic of computer operations, couched in terms of Turing Machines. In between come six chapters that attempt to span different uses of computers.

The material describing computers and programming is excellent. The author has done an unusual job in presenting a self-contained and understandable exposition, but without oversimplification or pedagogic tricks. This material can be read with profit either for a general picture of computers and programming methods; or, by carefully following the detailed examples and occasional exercises, the reader can go a good way into the actual art of computer programming. Thus the book can serve very well either as an auxiliary text or as the basis for a course in which the computer is used (for whatever reason).

In the introduction the author states that, in his discussion of the various uses

of computers, he presents examples with which he is most familiar, often work with which he has been involved. This gives clarity and authority in particular places, but it weakens the book considerably as a balanced and critical coverage of accomplishments in those fields using computers. Thus these chapters must really be considered as given examples of what can be done, as suggesting problems that the reader might be interested in pursuing, rather than as illuminating these problems or giving a balanced review of attacks that have been made.

These chapters are also extremely mixed in quality. The chapters on "Statistical Calculations and Data Analysis" and "Man-Computer Systems" are among the best. "Communicating in Natural Language: The Baseball Program," in general quite good, exhibits the defect of a distorted, narrow picture that inevitably results from giving a detailed discussion of only one of at least a half-dozen programs, all different in important ways. The chapter on "Computer Models of Psychological Processes" gives only poorly organized tidbits of what has been done here and there, with little illumination as to the direction or coherence of this work, or discussion that would suggest its important relevance to psychological theorizing. Again, the choice of material to be discussed should not be so casual in a book that is not a personal monograph, and purports to do more than exhibit random exercises. This seems especially unfortunate, since dynamic simulations of complex models are among the most fruitful, and least understood, uses for the computer. Two other extremely important uses, the control of on-line experiments by the computer, and hill-climbing techniques, are treated in two pages each. In contrast, an entire chapter is devoted to the relatively sterile and esoteric subject of random-number generators.

Despite these reservations about the balance, breadth and understanding exhibited in the content chapters, these chapters give a good enough picture of computer uses to whet almost any reader's appetite. And in fact they are probably better than most things the reader will find in book form today. Taken along with the sections on programming, information-processing and computer theory, they add up to what is probably the best textbook available today on computers in behavioral research.

LEONARD UHR
The University of Michigan

Studies in Item Analysis and Prediction. Edited by HERBERT SOLOMON. Stanford, California, Stanford University Press, 1961. Pp. ix, 310, $8.75.

The point has been made that psychological measuring instruments are similar to automobiles. The construction of both during the last forty years has not involved any new principles, although extensive development in the application of these principles has occurred in both fields. Readers of this book will no doubt wonder whether the time is not yet ripe for some new ideas in at least one of these fields.

Of the nineteen papers by ten writers collected together for this volume, at least six have appeared in published form previously. The papers are grouped into three parts. Part 1 consists of seven papers by Sitgreaves, Elfving, and Solomon in which most of the classical test theory assumptions are made: *e.g.* (1) a pool of test items is assumed to exist; (2) the items are assumed to be scored dichotomously, but associated with each item is a continuous normally distributed variable; (3) the criterion or unknown ability is also assumed to be normally distributed in a given population; and in some papers the factor structure of the criterion is assumed. Using this type of probability model and an occasional simplifying assumption when

the going gets rough, these papers attempt to determine optimum procedures for selecting items and estimating the unknown parameters associated with the criterion. The optimization is generally defined in terms of minimizing the mean squared error of estimate.

In Part 2, the seven papers by Lazarsfeld, Bahadur, Solomon, Raiffa, Johns, and Paulson attempt to dispense with the continuous variables associated with items and criterion. Except for Raiffa's chapter in which statistical theory is applied to the problems of item selection, the papers in this part seem mainly concerned with classification, the problem of deciding which of several populations an individual belongs to. An example cited in Solomon's chapter, of which there are very few throughout the book, consists of six science attitude items from which it is required to decide whether an individual has a high IQ or a low IQ.

In the third part, the papers by Sitgreaves, Teichroew, Elfving, and Bowker concentrate on the problem of obtaining the sampling distribution of Anderson's classification statistic.

Except perhaps for the last part, this book, according to the editor, is about psychological measurement and the design of psychological measuring instruments, yet it is very difficult to see any non-trivial psychological questions being asked or answered here. This is not to suggest that the book may not be both deep and interesting mathematically. One of the disturbing signs pointing to the emptiness of the psychological theorizing in this book is the fact that the work described could all too easily be applied to any scientific field whatever. It would not be necessary to rewrite a single chapter in this book if the problem posed was one of classifying stones instead of people.

A more serious discussion of psychological measurement theory and the design of psychological measuring instruments ought, it appears to this reviewer, to consider in some detail the nature of the pool of test items to be used—what kinds of behavior one should observe—and how the items are to be scored—what numerical representation theorem is to be established or can be established. The answer to both of these problems is presupposed here. What remains to be explored is outside the field of psychological measurement.

Indiana University — JOSEPH L. ZINNES

Emotions and Emotional Disorders: A Neurophysiological Study. By ERNST GELLHORN, and G. N. LOOFBOURROW. New York, Harper and Row, Publishers, Inc., 1963. Pp. xii, 496. $12.00.

This book is broken up into seven parts, the first covers the basic aspects of physiology necessary for comprehending general nervous system function. Naturally, the weakness of this section must lie in the tremendous amount of material it is necessary to cover. Basic references are, however, liberally given and where possible the pertinent work is described. From this first section, it is apparent that Gellhorn conceives of the hypothalamic system as the core to emotions. He attempts to integrate such diverse data as that emanating from the Magoun school, Papez, Grastyan, etc. in such a fashion as to arouse polemical thoughts for anyone with an investment in any structure other than the hypothalamus. Part II contains a major concept for Gellhorn, regulation of the autonomic nervous system and "tuning" of the hypothalamus. This section also gives a good description of the usage and general interpretation of the Mecholyl and Noradrenaline tests. Part III is that which

is likely to be most annoying to psychologists; for example, there might be quite an outcry against equating Pavlovian conditioning with "brain washing and so forth" (p. 177). At the same time, it should be noted that there are highly suggestive ideas in this section, such as the implication drawn by Gellhorn that positive conditioned stimuli act chiefly on the sympathetic system, while negative conditioned stimuli and the parasympathetic are related.

Parts IV and V are attempts to apply to experimental situations the concepts developed in earlier sections. The major concept constantly used is that of balance, and along with this the shifts from sympathetic to parasympathetic activity. In particular, these changes are applied to various psychosomatic disorders (hypertension, ulcer, asthma, etc.).

Part VI is probably the weakest of all, since it purports to deal with mental disorders but unfortunately, must reach too far to give the clinician any feeling of confidence or basically useful information. In Part VII, theories of emotion and future research are discussed.

This book would thus seem to have some limited usefulness as a source-book for an extensive list of references, particularly those relating to the hypothalamus, and a restricted but occasionally interesting overview of emotions as seen from the hypothalamus. It is also, perhaps, pertinent to note that the task of integrating the vast amounts of neuropharmacological and neurophysiological information available today is really formidable. With all the many limitations of this book it is a valuable effort at such an integration.

JOHN M. RHODES

Brain Research Institute
UCLA Medical Center
Los Angeles, California

Sleep and Wakefulness. By NATHANIEL KLEITMAN. Rev. ed. Chicago, University of Chicago Press, 1963. Pp. viii, 552. $12.50.

The original edition appeared in 1939. Though long out of print and still in demand, Dr. Kleitman resisted pressures of publishers merely to reprint. He saw the need for updating and extensive modification. Reflecting the voluminous activity in the last 25 years his new bibliography has been extended from 1434 to 4337 items, and the new text of only 370 pages cites most, if not all, of these references. This is possible since some pages list more than 100 references, *e.g.* p. 294, and numerous pages include at least 50 different items. Thus is illustrated the handbook or the encyclopedic nature of this volume.

The pattern of organization adheres to the original; in fact, the 36 chapter titles remain virtually unchanged. Dr. Kleitman, in the introduction, states that in arranging the material he has "endeavored to proceed from the observational through the experimental and pathological to the theoretical."

Since 1960 Emeritus Professor of Physiology at the University of Chicago, Dr. Kleitman now lives in Santa Monica, California. His retirement could not possibly represent any slackening of pace nor have his wide ranging interests diminished. He reports clearly on the psychological, pathological and pharmaceutical researches with the same vigor and decisiveness he displays on the more directly physiological content.

Readers will be most grateful for the author's referencing of works that have not been translated into English. Incidentally, titles in the bibliography have been trans-

lated unless they were written in English, French, German, Italian or Spanish. Each of the others shows parenthetically the language of publication.

Some of the recent Russian references are particularly interesting. For example, Kleitman quotes sections of Bogorad's collection in 1954 of articles supporting Pavlov's explanation of how spreading inhibition throughout the nervous system accounts for the onset of sleep. Kleitman aggressively tags the effort as "the extreme of Chauvinism" (p. 345), and as others have shown, this judgment may be justified. But the point is that the Kleitman offensive is evident throughout. It clearly influences selection and emphasizes works agreeing with Kleitman's earlier predilections. Contrary reports in many instances are given negative or cavalier treatment.

The conclusion is inescapable that this book, though encyclopedic and valuable manifests less humility and objectivity than the more classic first edition. Instead it is competitive and altogether too devoted to Kleitmanesque points of view. Unless the reader stays very fully awake he will be lulled into nodding agreement with unsupported agrument and convincing fallacy. This book can be recommended, therefore, only to the well warned and wary.

Virginia Military Institute — DEAN FOSTER

The Stream of Behavior. Edited by ROGER G. BARKER. New York, Meredith Publishing Company, 1963. Pp. x, 352. $6.75.

When someone gets around to writing a history of developmental psychology in a time to come, he will surely describe the present years as an era of controversy over research methodology. The controversy is basically whether the observation of behavior under naturally occurring conditions or the experimental manipulation of behavior under laboratory conditions is the fit and proper method of studying children. Roger Barker's commitment is firmly to the first approach, or more precisely—since observation is a plurality of techniques in itself—he is committed to the observation of relatively long, uninterrupted sequences of behavior.

The Stream of Behavior is a collection of field studies concerned with the "dynamic" processes of children's everyday lives and with the interactions between environmental forces and the child's behavior. (Adults are observed in one of the 10 field studies.) Preceding the field studies is a chapter by Barker devoted to those aspects of methodology that are common to the whole project. The second chapter, by Harold Dickman, is a methodological investigation of reliability in judging "natural" behavior units. The 10 field studies that follow are to some extent integrated, since the majority use as their basic data two or more of the 18 day-long observational records of normal and disabled children that have been collected under Barker's leadership. In each of these investigations, particular problems of methodology, together with the successes and failures in meeting these problems, are discussed at some length in conjunction with the presentation of method and results. The field studies cover these topics: Environmental forces in the everyday lives of children (Phil Schoggen); Mothers and fathers as sources of environmental pressure on children (Helen Simmons & Phil Schoggen); The social contacts of some midwest children with their parents and teachers (Arthur J. Dyck); Disturbances experienced by children in their natural habitats (Clifford L. Fawl); Social actions in the behavior streams of American and English children (Roger G. Barker & Louise Shedd Barker); Structure of the behavior of American and English children

(Maxine Schoggen, Louise Shedd Barker & Roger G. Barker); The behavior of the same child in different milieus (Paul V. Gump, Phil Schoggen & Fritz Redl); Some formal characteristics of the behavior of two disturbed boys (Nehemiah Jordon); A method of measuring the social weather of children (James E. Simpson); The study of spontaneous talk (William F. Soskin & Vera P. John).

University of Colo. Medical Center — ALMEDA ALLEN
University of Denver — YVONNE BRACKBILL

Language, Thought and Personality in Infancy and Childhood. By M. M. LEWIS. New York, Basic Books, 1963. Pp. 256. $5.00.

The central theme of Dr. Lewis' latest book is the role of language in cognitive and social-personal development from infancy onward. According to the author's thesis, language is a major factor in promoting the child's capacities for thinking and problem solving, in advancing the socialization process, and in promoting ethical behavior. Language and other aspects of behavior are interdependent rather than separate isolated strands of growth.

The material is divided into three parts corresponding to successive stages of child development—infancy, from birth to about three years, early childhood from three to seven, and later childhood from approximately seven to 12 years. Here as elsewhere the author follows Piaget's delineation of sequential growth periods.

Through the use of illustrations from his earlier studies on infant speech the author shows how linguistic development emerges through selective acts which result in the refinements which produce intelligible speech. Nearly half the book is devoted to this first period. In later sections the role of pre-adolescent sub-language in the processes of self-realization, of individuation and socialization are delineated through illustrations from numerous research studies. The author attaches great significance to the home environment, especially the intimate contacts of family life as factors in linguistic maturation, but he recognizes the paucity of research in this as in other areas. As a result he is forced at times to make conjectures based largely upon incidental information.

A unique feature of the book is the inclusion of material on the retarded development of deaf children and the socially deprived, especially institutionalized infants. Data concerning the adverse effects of language retardation on social behavior serve to illustrate the author's thesis and provide fuller understanding of linguistic growth processes in normal individuals.

Speech authorities and psychologists may not agree with the author's theories and interpretations of the known facts, but the existing evidence is presented in a well-organized manner and a readable style. Future studies will no doubt reveal more positively the influence of schooling and of the home environment on children's language. How does habitual TV viewing affect the child's linguistic and personal development? Here is a whole world of research in itself.

American University of Beirut, Lebanon — GERTRUDE HILDRETH

Theories in Contemporary Psychology. Edited by MELVIN H. MARX. New York, The Macmillan Company, 1963. Pp. xi, 628. $9.95.

In his preface, the editor sets forth the objective of the book as being "to encourage a more critical understanding and a sounder utilization of the principles of theory construction." Toward this objective, the editor has gathered together 35

articles, 23 of which were reprinted from various journals and books, and 12 of which were prepared especially for the volume. Basically, *Theories in Contemporary Psychology* is an updating and revision of an earlier compilation, *Psychological Theory*, which appeared in 1951.

The book is divided into three main parts: (1) theory construction, (2) special problems, and (3) fields of study. The first two parts concern themselves with general considerations of the role of theory in contemporary psychology and the problems and techniques of theory-construction. Models are emphasized over the grander theoretical systems. These first two sections include such topics as "The role of models," "Mathematical and statistical models," "Levels of explanation," "Theoretical constructs," and several others. The third part of the book considers particular theories in the various areas of psychology. This section encompasses theories from such areas as psychoanalysis, developmental psychology, social psychology, learning, sensation, and perception.

The scope of the book is surprisingly large and, being so, necessarily skips some areas and deals somewhat scantily with other areas.

The authority of the contributors is evident with such names as Stevens, Estes, Spence, Krech, Brunswik, Allport, Skinner, Leibowitz, and many others.

The book goes a respectible distance in fulfilling its objective. It is a good supplement to the handbooks which are generally outdated and unsatisfactory in their coverage of theoretical construction and the basic principles of models. The book is good, but could have been a great deal better had more schools of thought been represented in more than just passing comments.

Though basically designed for the graduate student, *Theories in Contemporary Psychology* would serve well as a source of readings for an undergraduate course in experimental psychology. It is recommended as a worthwhile addition to any library.

University of Texas — RAND B. EVANS

Understanding Interrelations in Nursing. By LESTER D. CROW and ALICE CROW. New York, The Macmillan Company, 1961. Pp. 461. $6.50.

Books on psychology for nurses run the risk of being too technical, too simple, too superficial, too comprehensive and too much like standard psychological texts. The current work suffers in most respects from all of these. It attempts to teach a basic psychology course at the high school level and to make it appropriate for the young nursing student. Eleven chapters, the core of the book, presents a sometimes simplified, sometimes complex general psychology. Except for questions asked at the end of each chapter, no attempt has been made to tie in what is presented with nursing training or experience. Six chapters attempt to integrate psychology and nursing but the effort is highly oversimplified. The title of the book is misleading, since "interrelations" are only a minor fraction of the content. Although probably a useful book for hospital-based nurse's training, it would need a skilled hand indeed to weave this traditional material into the needs of professional nursing.

Albert Einstein College of Medicine. — BERNARD KUTNER
Yeshiva University

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review.)

ADLER, ALFRED (eds. HEINZ L. ANSBACHER and ROWENA R. ANSBACHER), *Superiority and Social Interest. A Collection of Later Writings.* Evanston, Ill., Northwestern Univ. Press, 1964. Pp. xix, 432. $8.95.

ALCOCK, THEODORA, *The Rorschach in Practice.* Philadelphia, Lippincott, 1964. Pp. xii, 252. $12.50.

ALDER, HENRY L., and ROESSLER, EDWARD B., *Introduction to Probability and Statistics* (3rd ed.). San Francisco, Calif., W. H. Freeman & Co., 1964. Pp. xiv, 313. $5.50.

ANDREAS-SALOME, LOU (Trans by STANLEY A. LEAVY), *The Freud Journal of Lou Andreas-Salome.* New York, Basic Books, 1964. Pp. 211. $4.95 (paper).

ARON, RAYMOND, *German Sociology.* New York, Free Press of Glencoe, 1964. Pp. viii, 141. $1.95 (paper).

ATKINSON, JOHN W., *An Introduction to Motivation.* Princeton, N.J., Van Nostrand, 1964. Pp. xiii, 335. $9.75.

BAN, THOMAS, *Conditioning and Psychiatry.* Chicago, Ill., Aldine Publ. Co., 1964. Pp. ix, 244. $5.95.

BERGLER, EDMUND, *Parents Not Guilty! Of their Children's Neuroses. New Concepts in the Psychology of Human Behavior.* New York, Liveright, 1964. Pp. 283. $6.95.

BETTELHEIM, BRUNO, *Truants from Life: The Rehabilitation of Emotionally Disturbed Children.* New York, Free Press of Glencoe, 1964. (First published 1955.) Pp. xvi, 511. $2.95 (paper).

BRAZIER, MARY A. B., *Brain Function: Cortical Excitability and Steady Potentials; Relations of Basic Research to Space Biology.* Berkeley, Univ. of Calif. Press, 1964. Pp. xviii, 394. $10.00.

CARROLL, JOHN B., *Language and Thought.* Englewood Cliffs, N.J., Prentice-Hall, 1964. Pp. x, 118. Cloth, $3.95; Paper, $1.50.

CLECKLEY, HERVEY, *The Mask of Sanity* (4th ed.), St. Louis, Mo., C. V. Mosby Co., 1964. Pp. 510. $9.75.

CLINARD, MARSHALL B. (ed.), *Anomie and Deviant Behavior. A Discussion and Critique.* New York, Free Press of Glencoe, 1964. Pp. xii, 324. $6.50.

COHEN, ARTHUR R., *Attitude Change and Social Influence.* New York, Basic Books, 1964. Pp. xv, 156. $2.95.

CROW, LESTER D., and CROW, ALICE, *Adolescent Development and Adjustment* (2nd ed.), New York, McGraw-Hill, 1965. Pp. ix, 523. $7.95.

CROWNE, DOUGLAS P., and MARLOWE, DAVID, *The Approval Motive: Studies in Evaluative Dependence.* New York, Wiley, 1964. Pp. xii, 233. $7.50.

DAS, S. SUNDER, *General Psychology.* New York, Taplinger Publ. Co., 1964. Pp. xi, 220. $4.75.

DAVID, HENRY P., *International Resources in Clinical Psychology.* New York, McGraw-Hill, 1964. Pp. xiii, 236. $7.50.

DAVIES, E. BERESFORD (ed.), *Depression: A Cambridge Postgraduate Medical Course.* New York, Cambridge Univ. Press, 1964. Pp. xv, 378. $17.50.

DAVITZ, JOEL R., DAVITZ, LOIS, J., and LORGE, IRVING, *Terminology and Concepts in Mental Retardation.* New York, Bur. of Publ., Teachers Col., Columbia Univ., 1964. Pp. xiii, 127. $2.75 (paper).

DEMBER, WILLIAM, *Visual Perception: The Nineteenth Century.* New York, Wiley, 1964. Pp. xii, 222. $4.95.

DEWALD, PAUL A., *Psychotherapy: A Dynamic Approach.* New York, Basic Books, 1964. Pp. xvii, 307. $6.50.

DOOB, LEONARD W., *Patriotism and Nationalism: Their Psychological Foundations.* New Haven, Conn., Yale Univ. Press, 1964. Pp. xiii, 297. $6.75.

DREVER, JAMES, *A Dictionary of Psychology.* Baltimore, Md., Penguin Books, 1964. (First published 1952.) Pp. 316. $1.25 (paper).

EISENSTADT, S. N. (ed.), *Comparative Social Problems.* New York, Free Press of Glencoe, 1964. Pp. xi, 463. $9.95.

FLÜGEL, J. C., and WEST, DONALD J., *A Hundred Years of Psychology: 1833-1933; With Additional Part V: 1933-1963.* New York, Basic Books, 1964. Pp. 594. $6.00.

FODOR, NANDOR, *Between Two Worlds.* Englewood Cliffs, N.J., Prentice-Hall, 1964. Pp. xiii, 268. $7.50.

FORNELL, EARL WESLEY, *The Unhappy Medium. Spiritualism and the Life of Margaret Fox.* Austin, Tex., Univ. of Texas Press, 1964. Pp. x, 204. $5.00.

FRANKS, CYRIL M. (ed.), *Conditioning Techniques in Clinical Practice and Research.* New York, Springer Publ. Co., 1964. Pp. xii, 328. $8.50.

FREED, DONALD, *Freud and Stanislavsky. New Directions in the Performing Arts.* New York, Vantage Press, 1964. Pp. viii, 128. $2.75.

FROMM, ERICH (ed. by RUTH NANDA ANSHEN), *The Heart of Man, Its Genius for Good and Evil: Religious Perspectives.* New York, Harper & Row, 1964. Pp. 156. $3.95.

GEORGE, ALEXANDER L., and GEORGE, JULIETTE L., *Woodrow Wilson and Colonel House: A Personality Study.* New York, Dover Publ., 1964. (First published 1956.) Pp. xxii, 361. $2.00 (paper).

GINSBERG, MORRIS, *The Psychology of Society* (New ed.). New York, Barnes & Noble, 1964. Pp. xxiii, 187. $3.00.

GOODE, WILLIAM J., *The Family.* Englewood Cliffs, N.J., Prentice-Hall, 1964. Pp. viii, 120. Cloth, $3.95; Paper, $1.50.

GOODE, WILLIAM J. (ed.), *Readings on the Family and Society.* Englewood Cliffs, N.J., Prentice-Hall, 1964. Pp. xii, 242. $3.95 (paper).

GOULD, JULIUS, and KOLB, WILLIAM L., *A Dictionary of the Social Sciences.* New York, Free Press of Glencoe, 1964. Pp. xvi, 761. $19.50.

GREEN, MAURICE R., *Interpersonal Psychoanalysis: The Selected Papers of Clara M. Thompson.* New York, Basic Books, 1964. Pp. xii, 398. $8.50.

GROUP FOR THE ADVANCEMENT OF PSYCHIATRY, *Psychiatric Aspects of the Prevention of Nuclear War.* Rep. No. 57. New York, Group for the Advancement of Psychiatry, 1964. Pp. 223-317. Cloth, $3.00; Paper, $1.50.

GUENTHER, WILLIAM C., *Analysis of Variance.* Englewood Cliffs, N.J., Prentice-Hall, 1964. Pp. viii, 198. $6.95.

HADDON, WILLIAM, JR., SUCHMAN, EDWARD A., and KLEIN, DAVID, *Accident Research: Methods and Approaches.* New York, Harper & Row, 1964. Pp. xvi, 752. $15.00.

HAMMOND, PHILLIP E. (ed.), *Sociologists at Work: The Craft of Social Research.* New York, Basic Books, 1964. Pp. x, 401. $7.50.

HENRY, GEORGE W., *Masculinity and Femininity.* (Originally published as Parts I & II of *All the Sexes.*) New York, Collier Books, 1964. Pp. 320. 95¢ (paper).

HERING, EWALD (Trans. by L. M. HURVICH and D. JAMESON), *Outlines of a Theory of the Light Sense.* Cambridge, Mass., Harvard Univ. Press, 1964. (First published 1905, 1907, and 1911.) Pp. xxvii, 317. $9.95.

HOPKINS, TERENCE K., *The Exercise of Influence in Small Groups.* Totowa, N.J., Bedminster Press, 1964. Pp. viii, 205. $6.50.

HOUSSIADAS. LAMBROS, *An Exploratory Study of the Perception of Causality. Brit. J. Psychol. Monogr. Suppl.* XXXVI. New York, Cambridge Univ. Press, 1964. Pp. vi, 88. $4.50 (paper).

HOWES, RAYMOND F. (ed.), *Toward Better Preparation of College and University Administrators.* Washington, D.C., Assoc. for Higher Education, 1964. Pp. 37. $1.00 (paper).

HYMAN, RAY, *The Nature of Psychological Inquiry.* Englewood Cliffs, N.J., Prentice-Hall, 1964. Pp. xi, 116. Cloth, $3.95; Paper, $1.50.

JONES, EDWARD E., *Ingratiation: A Social Psychological Analysis.* New York, Appleton-Century-Crofts, 1964. Pp. ix, 211. $6.00.

KAHN, JACK H., and NURSTEN, JEAN P., *Unwillingly to School.* New York, Macmillan, 1964. Pp. xv, 187. $2.45.

KAHN, SAMUEL, *Psychodrama Explained*. New York, Philosophical Library, 1964. Pp. vi, 77. $3.00.
KASTENBAUM, ROBERT (ed.), *New Thoughts on Old Age*. New York, Springer Publ. Co., 1964. Pp. xii, 333. $7.50.
KATONA, GEORGE, *The Mass Consumption Society*. New York, McGraw-Hill, 1964. Pp. viii, 343. $7.95.
KIEV, ARI (ed.), *Magic, Faith, and Healing: Studies in Primitive Psychiatry Today*. New York, Free Press of Glencoe, 1964. Pp. xvii, 475. $6.95.
KNIGHT, JAMES A., *A Psychiatrist Looks at Religion and Health*. Nashville, Tenn., Abingdon Press, 1964. Pp. 207. $3.75.
KOESTLER, ARTHUR, *The Act of Creation: A Study of the Conscious and Unconscious Processes of Humor, Scientific Discovery, and Art*. New York, Macmillan, 1964. Pp. 751. $8.95.
KRONHAUSEN, PHYLLIS, and KRONHAUSEN, EBERHARD, *The Sexually Responsive Woman*. New York, Grove Press, 1964. Pp. 255. $5.95.
LANDIS, CARNEY (ed. by FRED A. METTLER), *Varieties of Psychopathological Experience*. New York, Holt, Rinehart & Winston, 1964. Pp. xviii, 484. $8.75.
LEIDERMAN, P. HERBERT, and SHAPIRO, DAVID (eds.), *Psychobiological Approaches to Social Behavior*. Stanford, Calif., Stanford Univ. Press, 1964. Pp. xv, 203. $6.75.
LENNEBERG, ERIC H. (ed.), *New Directions in the Study of Language*. Cambridge, Mass., The M.I.T. Press, 1964. Pp. ix, 194. $5.00.
LEVINSON, DANIEL J., and GALLAGHER, EUGENE B., *Patienthood in the Mental Hospital: An Analysis of Role, Personality, and Social Structure*. Boston, Houghton Mifflin, 1964. Pp. xvii, 265. $5.95.
LI, JEROME C. R., *Statistical Inference: I. A Non-Mathematical Exposition of the Theory of Statistics; II. The Multiple Regression and its Ramifications*. Ann Arbor, Mich., Edwards Bros., 1964. Pp. xix, 658; xiv, 575. $10.00 each.
MACE, C. A., *The Psychology of Study*. Baltimore, Md., Penguin Books, 1964. (First published 1932.) Pp. 119. 85¢ (paper).
MASLOW, ABRAHAM H., *Religions, Values, and Peak-Experiences*. Columbus, Ohio, Ohio State Univ. Press, 1964. Pp. xx, 123. $2.50.
MASSERMAN, JULES H., *Behavior and Neurosis: An Experimental Psychoanalytic Approach to Psychobiologic Principles*. New York, Hafner Publ. Co., 1964. (First published 1943.) Pp. xv, 301. $8.50.
MCGRATH, JOSEPH E., *Social Psychology: A Brief Introduction*. New York, Holt, Rinehart & Winston, 1964. Pp. ix, 166. $2.75 (paper).
MEERLOO, JOOST A. M., *Hidden Communion: Studies in the Communication Theory of Telepathy*. New York, Helix Press, 1964. Pp. 118. $4.75.
MIDONICK, HENRIETTA O., (ed.), *The Treasury of Mathematics: A Collection of Source Material in Mathematics Edited and Presented with Introductory Biographical and Historical Sketches*. New York, Philosophical Library, 1964. Pp. xxi, 820. $15.00.
MORGAN, CLIFFORD T., *Physiological Psychology* (3rd ed.). New York, McGraw-Hill, 1965. Pp. vii, 627. $8.95.
MURRAY, EDWARD J., *Motivation and Emotion*. Englewood Cliffs, N.J., Prentice-Hall, 1964. Pp. ix, 118. Cloth, $3.95; Paper, $1.50.
REDL, HELEN B., (ed. & trans.), *Soviet Educators on Soviet Education*. New York, Free Press of Glencoe, 1964. Pp. xxxvi, 252. $6.95.
REIK, THEODOR, *Voices from the Inaudible: The Patients Speak*. New York, Farrar, Straus & Co., 1964. Pp. 200. $4.50.
REISS, IRA L., *Premarital Sexual Standards in America*. New York, Free Press of Glencoe, 1964. Pp. 286. $1.95 (paper).
REISSMAN, FRANK, COHEN, JEROME, and PEARL, ARTHUR, *Mental Health of the Poor: New Treatment Approaches for Low Income People*. New York, Free Press of Glencoe, 1964. Pp. xv, 648. $9.95.
RHINE, J. B., *Extra-Sensory Perception*. Boston, Mass., Bruce Humphries Publishers, 1964. (First published 1934; reprinted with an additional introduction by the author). Pp. xlviii, 240. $2.75 (paper).

ROBACK, A. A., *A History of American Psychology.* (New, rev. ed.). New York, Collier Books, 1964. Pp. 575. $1.95 (paper).

SCHULTZ, DUANE P., *Panic Behavior: Discussion and Readings.* New York, Random House, 1964. Pp. viii, 146. $1.95 (paper).

SEBEOK, THOMAS A., HAYES, ALFRED S., and BATESON, MARY CATHERINE. *Approaches to Semiotics: Cultural Anthropology, Education, Linguistics, Psychiatry, Psychology. Transactions of the Indiana Univ. Conference on Paralinguistics and Kinesics.* Netherlands, The Hague, Mouton & Co., 1964. Pp. 294. Dutch Guilders 40.

SHELLY, MAYNARD, W., II, and BRYAN, GLENN (eds.), *Human Judgments and Optimality.* New York, Wiley, 1964. Pp. xiii, 436. $8.95.

SIMMEL, GEORG, *Conflict and The Web of Group-Affiliations: Two Major Essays on the Dynamics of Social Organization by the Great German Philosopher and Social Theorist.* New York, Free Press of Glencoe, 1964. Pp. 195. $1.95 (paper).

SOLOMON, DAVID (ed.), *LSD: The Consciousness-Expanding Drug.* New York, G. P. Putnam's Sons, 1964. Pp. xii, 273. $5.95.

STAATS, ARTHUR W. (ed.), *Human Learning: Studies Extending conditioning Principles to Complex Behavior.* New York, Holt, Rinehart & Winston, 1964. Pp. xv, 520. $8.75.

STEKEL, WILHELM, *Peculiarities of Behavior: A Perceptive Analysis of the Instincts and Emotions, Particularly of Sex.* Vols. 1 and 2. New York, Grove Press, 1964. (First published 1924.) Pp. xiv, 328, $1.75; Pp. x, 341, $1.75 (paper).

STEKEL, WILHELM, *Sexual Aberrations: Vol. 1. A Searching Analysis of the Sexual Distortion Known as Fetishism; Vol. 2. Fetishism in Sex: Onanism, Impotence, Sadism, Masochism, Transvestitism.* New York, Grove Press, 1964. (First published 1930.) Pp. ix, 369, $1.25; Pp. 355, $.75 (paper).

STENDLER, CELIA BURNS, *Readings in Child Behavior and Development* (2nd ed.). New York, Harcourt, Brace & World, 1964. Pp. xi, 498. $4.95 (paper).

STERN, PAUL J., *The Abnormal Person and His World: An Introduction to Psychopathology.* Princeton, N.J., Van Nostrand, 1964. Pp. xi, 239. $6.95.

STEVENS, HARVEY A., and HEBER, RICK (eds.), *Mental Retardation: A Review of Research.* Chicago, Univ. of Chicago Press, 1964. Pp. xii, 502. $12.50.

STORR, ANTHONY, *Sexual Deviation.* Baltimore, Md., Penguin Books, 1964. Pp. 139. 85¢ (paper).

STRASSER, STEPHAN, *Phänomenologie und Erfahrungswissenschaft vom Menschen.* Berlin, Walter de Gruyter & Co., 1964. Pp. xv, 313.

STRAUSS, ANSELM, SCHATZMAN, LEONARD, NUCHER, RUE, EHRLICH, DANUTA, and SABSHIN, MELVIN, *Psychiatric Ideologies and Institutions.* New York, Free Press of Glencoe, 1964. Pp. xi, 418. $8.50.

THOMSON, ROBERT, *The Psychology of Thinking.* Baltimore, Md., Penguin Books, 1964. (First published 1959.) Pp. 215. 95¢ (paper).

THYNE, JAMES M., *The Psychology of Learning and Techniques of Teaching.* New York, Philosophical Library, 1964. Pp. 240. $7.50.

TOFFLER, ALVIN, *The Culture Consumers: A Controversial Study of Culture and Affluence in America.* New York, St. Martin's Press, 1964. Pp. 263. $5.00.

TRAXEL, WERNER, *Einführung in die Methodik der Psychologie.* Bern and Stuttgart, Verlag Hans Huber, 1964. Pp. xi, 348.

VAN KREVELEN, D. ARN (ed.), *Child Psychiatry and Prevention.* Bern and Stuttgart, Verlag Hans Huber, 1964. Pp. 322. Fr./DM 36.

VON MENDELSSOHN, FELIX, *This is Psychiatry.* New York, Franklin Watts, 1964. Pp. xii, 237. $7.95.

VERNON, M. D., *The Psychology of Perception.* Baltimore, Md., Penguin Books, 1964. (First published 1962.) Pp. 265. $1.25 (paper).

WEITZ, HENRY, *Behavior Change through Guidance.* New York, Wiley, 1964. Pp. xiii, 225. $6.50.

WOLPE,, JOSEPH, DALTER, ANDREW, and REYNA, L. J., *The Conditioning Therapies. The Challenge in Psychotherapy.* New York, Holt, Rinehart & Winston, 1964. Pp. viii, 192. $7.00.

ARE YOU MOVING?

If you advise the Post Office of your willingness to pay the postage, your copies of the JOURNAL will be forwarded to you. If you do not, you will not receive them—
UNLESS
you advise this Office of your change of address before the quarterly numbers are mailed.

The latter is by far the better plan.

HEREAFTER—

Since, under new rules and regulations, the Post Office no longer returns undeliverable copies to the JOURNAL's Office (the *label*, indicating change of address, is *sometimes* sent to the publisher), there will be a single-copy charge of $2.00 for any issue of the JOURNAL sent from this Office because of your failure to notify us of a change in your address in advance of publication.

LET US KNOW—

Please, therefore, notify us not later than the first day of the month of publication —March, June, September, December—of a change in your address, and please give us the "old" address as well as the "new" one. We are interested in getting your copies of the JOURNAL to you.

Business Office

THE AMERICAN JOURNAL OF PSYCHOLOGY

Mezes Hall 102, University of Texas

Austin, Texas 78712

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

VOL. LXXVIII JUNE, 1965 No. 2

PATTERN-DETECTION IN BRAIN-INJURED CHILDREN

By CHARLES F. REED and ALAN POLLACK,
Temple University and University of California at Los Angeles

In traditional tests using hidden figures, the *S* is required to identify line-drawn figures variously camouflaged by other lines. He may be asked to name or point out disguised common objects or, as in the classical Gottschaldt figures, to trace the outlines of a geometric figure concealed in a complex design. The figures may be tachistosocopically presented.[1]

In experiments with normal *Ss*, performance has been reported to vary with age and, less consistently, with intelligence and sex.[2] Damage to the central nervous system (as in the case of soldiers with penetrating missile wounds to the head, brain-injured children, and cerebral palsied children) has been found consistently to be associated with poor performance.[3] The difficulty for brain-injured *Ss* has

* Received for publication December 4, 1961. This research was supported by USPHS Grant M-4761 (A) and, 2M-5929-C6. The General Electric Company assisted in the preparation of displays, and the Walter D. Matheny School, the A. Harry Moore School, and the Syracuse Public Schools provided facilities and subjects.

[1] W. M. Cruickshank, H. V. Bice, and N. E. Wallen, *Perception and Cerebral Palsy*, 1957, 37-46; Lila Ghent, Perception of overlapping and embedded figures by children of different ages, this JOURNAL 69, 1956, 575-587; Kurt Gottschaldt, Ueber den einfluss der Erfahrung auf die Wahrnehmung von Figuren, *Psychol. Forsch.*, 8, 1926, 261-317; 12, 1929, 1-87; L. L. Thurstone, *A Factorial Study of Perception*, 1944, 72; H. A. Witkin, Individual differences in ease of perception of embedded figures, *J. Pers.*, 19, 1950, 1-15; Heinz Werner and A. A. Strauss, Pathology of figure-background relation in the child, *J. abnorm. soc. Psychol.*, 36, 1941, 236-248.

[2] Leonard Cobrinik, The performance of brain-injured children on hidden-figure tasks, this JOURNAL, 72, 1959, 566-571; Ghent, *op. cit.*, 587; H-L. Teuber, Some alterations in behavior after cerebral lesions in man, in Amer. Assoc. Adv. Sci., *Evolution in Nervous Control*, 1959; Cruickshank, Bice, and Wallen, *op. cit.*, 105. See also summaries of Teuber, Perception, in H. W. Magoun (ed.), *Neurophysiology*, 3, 1960, 1609, and in Witkin, *et al.*, *Personality through Perception*, 1954, 477-478.

[3] Walther Poppelreuter, *Die psychischen Schädigungen durch Kopfschuss im*

been attributed to distractability,[4] inability to organize individual stimuli into wholes,[5] and 'pathology' in the figure-background orientation of the visual field.[6] Teuber considers the phenomenon to be a non-specific, non-localizable consequence of brain-injury. He and his associates found poor performance to be independent of the site of the lesion and of the presence of symptomatology, though aphasic patients showed marked deficiency.[7]

Several aspects of the traditional testing procedures limit sensitive measurement and, especially in the case of brain-injured *Ss*, may contaminate the experimental findings.

Motor handicap, for example, probably affects performance beyond the execution of the action of pointing out or tracing figures. Cobrinik found severity of motor impairment to be significantly related to the test-performance of cerebral palsied children, though he does not report whether such impairment was obvious in the act of pointing.[8] It anticipates some of the results of the present study to suggest that control of eye-movements may limit the examination of test-figures. Tachistoscopic studies have used exposure-times which did not preclude the possibility of eye-movement. Not only do children require longer tachistoscopic exposure for the perception of pattern than do adults, but brain lesions in children produce additional slowness in visual searching of patterns.[9] Moreover, any score based upon time of recognition assumes equivalent familiarity of the test-figures for all *Ss*. This assumption may be unwarranted in the case of children with cerebral palsy who, as a consequence of motor handicap, probably have approached, touched, and handled objects with less frequency and facility than have normal children.

A second questionable assumption concerns the adequacy of the communicative capacities of *S*. In studies using hidden figures, internal check of the assumption has been lacking. *S*'s literal response to the instruction "to tell what he sees" has constituted the basis for differential scoring, *e.g.* on whether, in the judgment of *E*, an object was precisely or vaguely described.[10] Equivalent communicative capacities are especially unlikely when the *Ss* include both brain-injured and normal children.

For the most part, the displays using hidden-figures have been ad hoc constructions. They have not been ordered in difficulty on the basis of defined structural characteristics. Hence it has not been possible to specify why displays have differed in difficulty or to prepare tests of equivalent or graded difficulty. There have been

---

Kriege 1914-16, 1917, 189-190; Werner and Strauss, *op. cit.*, 236-248; H-L. Teuber, W. S. Battersby, and M. B. Bender, Performance of complex visual tasks after cerebral lesions, *J. nerv. ment. Dis.*, 114, 1951, 413-429; H-L. Teuber and Sidney Weinstein, Ability to discover hidden figures after cerebral lesions, *A.M.A. Arch. Neurol. Psychiat.*, 76, 1956, 369-377; W. M. Cruickshank and J. E. Dolphin, Figure-background relationship in children with cerebral palsy, *J. clin. Psychol.*, 7, 1951, 228-231; Cobrinik, *op. cit.*, 566-571.

[4] Werner and Strauss, *op. cit.*, 247-248.

[5] Cruickshank, Bice, and Wallen, *op. cit.*, 114.

[6] Werner and Strauss, *op. cit.*, 245-248.

[7] Teuber and Weinstein, *op. cit.*, 369-377.

[8] Cobrinik, *op. cit.*, 571.

[9] Teuber, *op. cit.*, 1960, 1610.

[10] Cruickshank, Bice, and Wallen, *op. cit.*, 1957, 44-45; Cruickshank and Dolphin, *op. cit.*, 229; Werner and Strauss, *op. cit.*, 237-238.

attempts, however, to identify significant variables in the visual displays. Ghent has observed that figures sharing contours may be more difficult to isolate from each other than are figures with intersecting contours;[11] data consistent with the interpretation have been gathered.[12] Similarly, Osgood has noticed that the simple shapes in the Gottschaldt figures were obscured best by lines extending their contours and presumably inducing erroneous eye-movements.[13]

Green, Wolf, and White have studied perception of form as a problem in the detection of a 'signal' in a 'noisy' display. They prepared visual displays in which noise was defined in terms of the statistical properties of arrays of dots.[14] The task appears logically related to tests employing hidden-figures. Hidden-figures may be viewed as complex forms in which a given form is to be isolated from interfering, more-or-less 'noisy' fields; in the traditional figures, interference is not random but organized into competing forms. Green, Wolf, and White have demonstrated that the curve of detection in their displays for normal adult *S*s resembled, with higher thresholds, the curve of detection of an hypothetical noise-free detector operating on the statistical characteristics of the displays. It may be conjectured that an *S* with damage to his central nervous system would have even higher thresholds.

The objective of the present experiment was to test pattern-detection in brain-injured *S*s. Certain perceptual consequences of brain-injured *S*s have been delineated in the experiments using hidden-figures. The technique of utilizing statistical displays appears to provide a method for improved measurement of those perceptual effects through quantification of defined interference and reduction of contaminants.

Method: (1) Materials and procedures. The preparation of visual displays followed the procedures of Green, Wolf, and White, but with modifications made necessary by the facilities available.

Green photographed matrices of dots as they appeared on an oscilloscope attached to a digital computer. A square matrix was subdivided into a set of equal bars with the positions in the odd-numbered bars filled with dots in accordance with an assigned probability p_1; the positions in the even-numbered bars were filled in accordance with a probability p_2.

In the present experiment, the output of the computer was converted into printouts on white paper. The printed dots were, however, slightly higher than they were wide. To produce a square display, it was necessary to include a larger number of spaces in the horizontal dimension of the display than in the vertical (Fig. I). All displays were squared in this fashion, then photographed with bars oriented both horizontally and vertically. The potential clue to orientation which existed in the shape of the individual dots was not apparent to adult *S*s tested with the displays at the distance used in the experiment.

[11] Ghent, *op. cit.*, 576-587.
[12] Cobrinik, *op. cit.*, 566-571.
[13] C. E. Osgood, *Method and Theory in Experimental Psychology*, 1953, 217.
[14] B. F. Green, Jr., A. K. Wolf, and B. W. White, The detection of statistically defined patterns in a matrix of dots, this JOURNAL, 72, 1959, 503-520.

In all displays in the present experiment, p_2 was made equal to $(1 - p_1)$. Displays were prepared with values of p_1 from 0.90 to 0.53. The differences in probability (delta p) between adjacent bars ranged from 0.80 to 0.06. Each display consisted of a six-bar matrix.[15]

In the experiment, displays were so projected that they filled completely a screen measuring 11.5 × 11.5 in. S was seated beside the projector in a darkened

FIG. 1. SAMPLE OF THE VISUAL DISPLAY USED

room, 5.8 ft. from the screen's position on blank wall. The visual angle subtended by the displays was about 9.5°, varying slightly because of freedom of head-movement.

The same series of displays was utilized for all Ss. The order of presentation was as follows (in terms of p_1 value and Vertical (V) or Horizontal (H) orientation): 0.90 V, 0.80 H, 0.70 H, 0.80 V, 0.90 H, 0.70 V, 0.60 V, 0.65 V, 0.54 V, 0.90 H, 0.56 V, 0.65 H, 0.60 H, 0.56 H, 0.59 V, 0.58 V, 0.55 H, 0.80 V, 0.54 H, 0.57 V, 0.70 H, 0.59 H, 0.55 V, 0.58 H, 0.57 H, 0.53 V, 0.53 H. The first six displays were used to instruct the S and to provide an indication of S's understanding of the task. If a child was unable to report the orientation of all six displays, he was not tested further. Six cerebral palsied children out of 105 tested were

[15] Decimal points are omitted from delta-p values in the remaining discussion.

excluded on this basis. Three added instances of the easy displays of the practice series were scattered in the test-series to provide experiences of success and recall of the task.

Each display was shown until *S* reported whether he saw horizontal or vertical bars. Most *Ss* answered within 5.0 sec., but they were encouraged to guess when they experienced difficulty. Since the *Ss* were children of a wide range of ages, some with speech and motor handicap, the instructions were adapted to the understanding and communicative capacities of each child. In some instances communication was non-verbal, *e.g.* the child indicated response to a display by pointing toward one of two printed sample displays put at his side. Again, the criterion for successful communication and understanding was an accurate report on the six displays of the practice series.

Subjects. The brain-injured *Ss* in the experiments were cerebral palsied children obtained from one private and two public schools for handicapped children. No attempt was made to select the children on the basis of intelligence test-scores. At the private and at one public school all children capable of attending to the task and communicating a choice by some means were examined. Inclusion in the sample depended upon successful responses to the practice displays. At the second public school, a sample of *Ss* was taken, again depending only upon ability to respond to the practice displays.

The combined groups of cerebral palsied children were distributed by diagnostic categories as follows: spastic, 54; athetoid, 29; ataxic, 5; rigid, 5; mixed, 4. Two additional *Ss* had been given designations of "brain-injury in childhood" without further description. Ten of the children were not included in the analysis of results because it was not possible to obtain visual refractions on them or because they were designated "legally blind," or "having very poor vision." Degree of physical handicap varied from very mild, to very severe, including speech and hearing defects. The 50 boys and 30 girls were equally distributed over the diagnostic categories.

Children under 7 yr. of age were not tested; the total sample ranged in age from 7 to 19 yr.

The control *Ss* were students without handicap at two public schools. The sample consisted of 36 boys and 34 girls, drawn to match the age-distribution of the experimental sample, again without regard to intelligence test-scores.[16]

Additional variables. The refractive characteristics of each *S*'s eyes were ascertained by the method of static retinoscopy.[17] Light reflected from the retina is used to determine the error of refraction of the principal axes of the eye. No report is required from *S*. The *Ss* who wore glasses were tested while wearing them; the object of the examination was to measure the uncorrected residual error that might influence the examination of the displays. Refractive measures are ex-

---

[16] At the conclusion of testing for the original sample of cerebral palsied children, the display slides were damaged and had to be replaced. To insure that the new slides, which appeared somewhat darker, were equivalent to the old slides, it was necessary to enlarge the cerebral palsied sample and compare results with the original sample. The scores were not different and it was concluded that the displays were equivalent.

[17] S. Duke-Elder, *The Practice of Refraction*, 1954, 202-220.

pressed in terms of the power of lenses required to compensate for the refractive condition of *S*'s eyes—negative, to correct myopia, and positive, to correct hypermetropia.

Children with other ocular defects, either apparent or noted in available clinical records, were treated separately in the statistical analysis. Ocular defects included strabismus, large muscular imbalance, end-point nystagmus, and difficulty in visual pursuit.

The *S*s were not matched on the basis of intelligence test-scores. Apart from the difficulty of equating motor-handicapped children with non-handicapped children, the visual task seems to require a perceptual achievement prior to any of the tasks presented in intelligence tests. Nevertheless, the relationship of 'intelligence' to the experimental variable was evaluated by means of the vocabulary scale of Form L of the Stanford-Binet. Scores were expressed in terms of the number of words defined.[18]

Results. Performance on the perceptual task was translated into two expressions: 'Total Error' and 'Initial Error.' 'Total Error' refers to the sum of errors made by an *S* on the 21 displays of the test-series, regardless of the delta-*p* level at which the error occurred. 'Initial Error' is a score which assigns to *S*'s performance the numerical value of the 'easiest' display failed, *i.e.* the display with the highest delta-*p* value. The score ignores the number of errors made; an *S* making a single error at a relatively high delta-*p* value receives a score implying poor performance.

TABLE I

Total Number of Errors on 21 Tests and delta-*p* Value of initial Error (Cerebral palsied and normal children.)

		Total errors		Initial error	
S	*N*	Mean	*SD*	Mean	*SD*
Normal	70	3.06	2.02	12.60	5.53
Cerebral palsied:					
Spastic	29	6.45	2.77	21.31	9.43
Athetoid	14	6.71	2.84	23.14	17.61
Miscellaneous	12	7.25	2.31	21.17	7.48
With ocular defects	34	7.68	1.92	22.59	7.07

Table I lists the mean results for the normal group and for the four sub-groups of brain-injured children.[19] For both 'Total Error' and 'Initial Error' scores, an analysis of variance established that the difference between the normal children and the several groups of brain-injured children was significant beyond the 0.001 level of significance. The children in the diagnostic sub-categories of cerebral palsy do not differ in either

[18] L. M. Terman and M. A. Merrill, *Measuring Intelligence*, 1937, 302-322.

[19] Correlation between Total Error and Initial Error scores was 0.80 for the total sample, 0.63 for normal sample.

performance-measure, but in view of the small number in each category, only a tentative inference is possible.

The mean scores of palsied children with ocular defects is not significantly inferior to the means of palsied children without apparent defect. In the case of the Initial Error scores, however, there are differences in the variability within the sub-groups which require examination.

Accordingly, Chi-squared analysis was performed on the frequency-distributions of Table II. As in Table I, the normal children perform

TABLE II

FREQUENCY-DISTRIBUTION OF INITIAL ERRORS

Delta-p	Normal	Cerebral palsied without ocular defect			Cerebral palsied with ocular defect
		spastic	athetoid	miscellaneous	
.40+	0	4	1	1	2
.30	2	3	0	1	11
.20	10	9	5	7	11
.18	4	5	5	1	3
.16	6	0	1	0	5
.14	6	5	0	2	1
.12	15	1	1	0	0
.10	0	0	0	0	1
.08	18	2	1	0	0
.06	9	0	0	0	0
Total	70	29	14	12	34

better than the cerebral palsied children, who, at the same time, do not differ by diagnostic sub-category. The cerebral palsied *Ss* with ocular defect are distinguished, however, from the cerebral palsied children without ocular defect ($p < 0.01$). Whether this result is a peculiarity of the Initial Error score or an indication of a real difference in sensitivity remains to be determined. Fig. 2 portrays the cumulative distributions on the Initial Error score.

Additional measures. Table III summarizes age, vocabulary-scores and refractive measures for the various groups.

The refractive measurements are expressed in eight indices. The eye which yielded the smallest refractive error in any meridian was called the 'better' eye. The greater and lesser curvatures of both eyes formed the first four indices; all of these measures ignore astigmatic conditions. The second four measures attempted to deal with astigmia; they express the differences in the curvature of the chief meridians in each eye. Measures VII and VIII were computed to allow for the possibility that reliance upon an eye is a function of degree of astigmatism; VII represents the eye with the smallest difference between meridians; VIII, the larger differences between meridians.

The entries for the first four refractive measures in Table III reflect an over-all uncorrected myopia. This is a result to be expected in the absence of cycloplegic drugs.

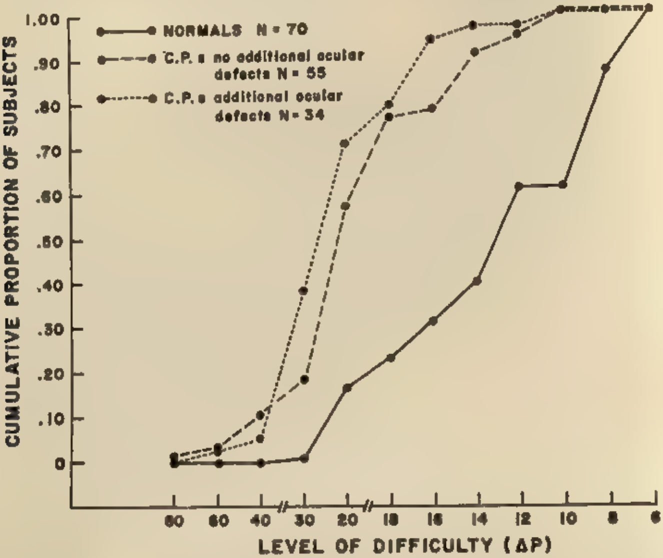

Fig. 2. Cumulative Distributions of the Initial Errors of the Normal and the Cerebral Palsied *Ss*

TABLE III

Mean Refractive Indices, Age, and Vocabulary-Scores for Cerebral Palsied and Control Groups Tested with Randomized Dot-Patterns

Residual refractive measures (in diopters)	Normal *Ss*		Cerebral palsied *Ss* without ocular defect: spastic		athetoid		misc.		Cerebral palsied *Ss* with ocular defect	
	mean	*SD*	mean	*SD*	mean	*SD*	mean	*SD*	mean	*SD*
I*	−.29	.66	−.44	.55	−.28	1.14	−.26	.60	−.27	1.07
II	−.31	.40	−.19	.26	−.14	.63	−.19	.37	−.05	.75
III	−.65	.71	−.81	.70	−.09	1.44	−.77	.66	−.73	1.65
IV†	−.55	.61	−.58	.68	−.27	.98	−.29	.59	−.52	1.20
V	.26	.40	.43	.51	.50	.48	.43	.48	.33	.61
VI	.23	.38	.32	.43	.50	.57	.52	.48	.51	.79
VII	.12	.31	.21	.28	.32	.47	.33	.42	.21	.56
VIII	.36	.45	.56	.61	.68	.52	.62	.50	.61	.78
Age (in yr.)	13.48	2.75	13.10	3.16	13.05	2.69	12.46	3.40	14.33	3.67
Vocabulary	18.26	4.70	15.04	5.65	14.58	4.84	11.88	5.60	15.33	6.00
	70		29		14		12		34	

* I. Greater curvature of better eye; II. Lesser curvature of better eye; III. Greater curvature of poorer eye; IV. Lesser curvature of poorer eye; V. Cylindrical difference in better eye; VI. Cylindrical difference in poorer eye; VII. Lesser cylindrical difference; and VIII. Greater cylinidrical difference.

† Measures I–IV represent power of lenses needed to bring refraction of *S*'s eye to neutral point.

The correlations of vocabularly and refractive scores with the performance scores are compiled in Table IV. They are within-group correlations taken from analyses of covariance in which each of the screening measures was treated as an uncontrolled

variable. The negative correlations between Total Error and the first four refractive measures indicates an association between poor performance and myopia. By contrast, the Initial Error scores do not show a similar tendency; the correlations for refractive measures III and IV in fact indicate an opposite tendency. It seems possible to rationalize these correlations by supposing that the displays with small differences between bars (small delta-p) were disproportionately difficult for the relatively myopic *Ss*. It may have been possible for a small degree of uncorrected myopia to be of assistance at higher delta-p levels, where blurring would assist contour-formation.

Vocabulary-scores show a slight inverse relationship to error; the correlations are significant at the 0.01 level.

Analysis of covariance for the error-scores, considering each of the refractive indices and the Vocabulary-Score in turn as uncontrolled variables, indicated that

TABLE IV

CORRELATIONS OF REFRACTION AND VOCABULARY VARIABLES WITH PERFORMANCE-MEASURES

Screening Measure		Total error	Initial error
Refraction	I	−.355*	−.098
	II	−.236*	.167
	III	−.234*	.381*
	IV	−.250*	.484*
	V	.017	.106
	VI	.128	.361*
	VII	.043	.020
	VIII	.089	−.034
Vocabulary		−.239*	−.231*

* 0.01 level of significance

the difference between the normal children and the cerebral palsied children could not be attributed to differences in these variables. All F-ratios were significant beyond the 0.001 level of significance.

Further analysis of other variables revealed no relationship with performance. The horizontal and vertical orientations of the displays posed no differential difficulty. Sex was unrelated to performance and, within the range tested, age was unrelated to performance.

Discussion. The findings of the experiment are in accord with the conjecture that brain-injured *Ss* are less able to perceive simple forms in 'noisy' displays than are normal *Ss*. The deficit does not depend upon differences in refractive powers of the eye. It does not depend upon differences in intelligence, as far as that complex variable is represented in achievement of a vocabulary. Detection of patterns in the displays is independent of sex and, within the range tested, of age.

There is some evidence that cerebral palsied children with ocular defects perform even less adequately than do those without apparent defects. When performance is measured on the basis of the easiest display failed, the children with identified defects are the less able. The

present experiment used only a relatively small number of displays in a fixed series. This limitation leaves unsettled which of the scores used more adequately reflects performance. The possibility of guessing correctly enters particularly into the Total Error score. Momentary distractions and fluctuations of attention, on the other hand, would be penalized by the Initial Error score. Only an enlarged series of displays and repeated measurement will make clear the relationship of ocular defect to performance.

The results of the present experiment do suggest that control of the movements of the eyes may be essential to pattern-detection and to performance on the hidden-figure tasks. It is, of course, possible to hypothesize a basis for poor performance in the visual pathway and in central information-processing functions. In general, experiments with hidden-figures have assumed that central events were responsible for observed differences among *S*s. While this may be the case, these experiments have not examined visual functions, beyond excluding *S*s with obvious visual defects. In the present experiment, it is conceivable that children suffering cerebral palsy without apparent ocular defects have, in fact, subtle oculomotor handicaps which limit their examination of the visual displays. The abilities to scan systematically, to follow contours, or to sustain fixation, may all be required for detection of patterns in our displays.

An alternative possibility is that the cerebral palsied children with ocular defects are in some sense more severely brain-damanged than are the children without ocular defects. The pattern-detection task would be simply a reflection of extent of damage. Presumably, if this were the case, the children with ocular defect would show poor performance in other cognitive tasks. The results of the Vocabulary examination do not support this possibility, but they also do not exclude it.

The foregoing discussion has been concerned with variables within the *S*s. Two questions concerning the nature of the task remain.

It was argued that the traditional task of finding a hidden-figure could be reduced to the simpler pattern-detection of randomized displays defined in terms of their statistical properties. That equivalence is not demonstrated easily, especially if the test of the hidden-figure incorporates the complications which were attributed to it. If the statistical displays represent the essential perceptual problem, poor performance on them would imply poor performance on the traditional figures, though not necessarily the converse.

A second question awaiting resolution is the possibility that the statistical displays are simply tests of brightness-discrimination. Green, Wolf, and White, considering this problem, observed that the threshold for discrimination of uniform

fields was different from the threshold for the detection of the barred pattern, when the latter was converted to a brightness-difference.[30] A more direct comparison of performance on dot-displays and on uniform fields is required.

Summary

The traditional hidden-figure tests may impose unusual requirements when used to test brain-injured *S*s. For this reason, the task was redefined in this study as the detection of a pattern of bars in a 'noisy' visual display. The displays were generated by a technique described by Green, Wolf, and White. The task for 89 cerebral palsied children and 70 normal children was to report the horizontal or vertical orientation of bars formed by arrays of dots. Distributions of age (7 to 19 yr.) and sex were similar for both groups of *S*s. Uncorrected refractive errors were measured by the method of static retinoscopy, and the Stanford-Binet Form L Vocabulary scale was administered.

The performance of the cerebral palsied children was inferior to that of the normal children; the difference in performance could not be attributed to differences in the refractive and intelligence-measures. Some evidences was found that cerebral palsied children with ocular defects (muscular imbalances, end-point nystagmus, defects in visual pursuit) differed in performance from those without ocular defects. Oculomotor control may be a significant factor in the performance of brain-injured *S*s in pattern-detection and in the traditional hidden-figure tests; further investigation of this point is required.

[30] Green, Wolf, White, *op. cit.*, 519-520. In the present experiment, the ratio of black area to total area in each bar is 0.492.

SEEKING INFORMATION TO REDUCE THE RISK OF DECISIONS

By Ward Edwards and Paul Slovic, University of Michigan

Many, perhaps most, decisions are concerned with the acquisition of information relevant to some final or terminal decision.[1] Information usually is costly, and the extent of its relevance to a terminal decision seldom is guaranteed in advance. The considerations of cost, pay-off, and probability that apply to ultimate decisions also should be applied to decisions whether or not to seek information. The appropriate criterion is easy to state, although often hard to apply. Information should be sought only if the expected cost of obtaining it is less than the expected gain from it.

Ideal performance in the simplest multi-decision, information-seeking problems can be specified with no tools other than the notion of expected value, and this paper reports an experiment based on such a simple problem. The *S*s made a sequence of decisions about whether or not to buy information which would reduce or eliminate the risk incurred in making a final decision. The task somewhat resembled that of a military commander who must decide between committing troops or weapons on the basis of sketchy, incomplete information, or instead, accepting the cost (in time, men, materiel, or all three) of seeking additional information which may or may not help him much.

Method

Subjects. Ten undergraduate men participated in 11 experimental sessions. In each session, each *S* performed 48 tasks. The costs and pay-offs of the tasks varied from session to session, but were the same for all the tasks in any given session.

Standard task. *S* was confronted with 48 16-cell square matrices and told that he would be paid a specified amount of money for finding the unique cell in each matrix. He was allowed one free guess. Before making it, he could buy the oppor-

* Received for publication April 16, 1964. This work was sponsored in part by Project MICHIGAN under Department of the Army Contract DA-36-039 SC-78801 and in part by AF AFOSR 192-63. The authors are grateful to Robert Ollman for assistance in taking and processing data; to Harold Lindman for writing a computer-program; and to Mrs. Carol Orwant, who invented the punchboard, helped design the study, took data, and did much of the data-processing. Paul Slovic is now at Oregon Research Institute, Eugene, Oregon.

[1] Ward Edwards, Dynamic decision theory and probabilistic information processing, *Human Factors*, 4, 1962, 59-73.

tunity to look at a cell to find out whether or not it was the unique cell. After that look, he could buy another look, and so on up to 15 looks for each matrix. After finding the unique cell or taking his free guess, S went on to the next matrix. The unique cells were located equiprobably.

All the strategies used by Ss can be described by the rule: Look until you find the unique cell or until you have taken L looks, whichever comes first; if after L looks, you have not found the unique cell, guess. The 16 such strategies, defined by $L = 0$ through $L = 15$, were numbered 1-16, corresponding to the maximal number of cells which might be looked at under each strategy; because of the free guess, that maximal number always was one more than the maximal number of looks which might be bought under the strategy. If N is the number of the strategy, R is the reward for finding the unique cell, and C_n is the cost of the nth purchased look, then the expected value of the Nth strategy is:

$$EV_N = N/16(R) - \sum_{n=1}^{n=N-1} [(17-n)/16][C_n]$$

Table I gives the values of R and C_n for the various conditions of the experiment, and Fig. 1 gives plots of EV_N as a function of N as well as the response-data. Each task is identified by the number of the session devoted to it. The optimal strategy for any task is easily defined: choose that value of N for which EV_N is greatest.

Of the four standard tasks, Tasks 1 and 6 were alike in that the cost of information was fixed at 2¢ per look. In the other two standard tasks, 4 and 9, the cost of looking varied.

Inverse task. A task formally very similar but psychologically quite different was made simply by reversing the signs of the costs and payoffs for the standard task. In this inverse task, S was paid for each cell he looked at, except that if he found the unique cell, he was fined.

To make the formal similarity between the two tasks as close as possible, we required S to investigate at least one cell; he was fined the specified amount if he found the unique cell but was not paid for that first look. Unless he found the unique cell, S was paid a specified reward for each look after the first, if he chose to take any. If on any look he found the unique cell, he was not rewarded for that look, but fined. As in the standard task, if N is the maximal number of looks (counting the required first look) that S contemplates taking, C is the fine charged for finding the unique cell, and R_n is the reward obtained by taking the Nth look (not counting the required first look, and provided that S does not find the unique cell on the Nth look), then the expected value of the Nth strategy is:

$$EV_N = (-N/16)C + \sum_{n=1}^{n=N-1} [(17-n)/16)][R_n]$$

Apparatus. The only apparatus used was a simple punchboard. A backing plate, a sheet of paper, and a transparent plastic sheet were assembled in that order in a holder. The plastic sheet on top had 204 round holes in it, each ⅜-in. in diameter, arranged in 12 groups of 17 holes each. Within each group, 16 holes were arranged in a square matrix; the 17th hole, where S wrote down his free guess, was located below the lower right-hand hole of the matrix. This pattern was re-

produced on the paper underneath the plastic sheet. On the backing plate, behind the paper, there was one hole for the unique cell in each of the 12 matrices. *S* investigated a hole by writing in it with a sharp hard pencil; he wrote *A, B, C, . . .* for his information-seeking moves, and *X* for his free guesses. If he tried to write in

TABLE I

COSTS AND PAY-OFFS

		Session number				
		1	2	6	7	8
Tasks for which cell-cost or pay-off was fixed	Task-type	standard	mirror	standard	inverse	inverse
	Pay-off or cost in cents for finding unique cell	+20	−20	+12	−12	−12
	Cost or pay-off in cents for each voluntary look	−2	+2	−2	+2	+1

			Task number					
			3	4	5	9	10	11
Tasks for which cell-cost or pay-off varied	Task-type:		inverse	standard	inverse	standard	inverse	inverse
	Pay-off or cost in cents for finding unique cell		−20	+32	−32	+40	−40	−40
	Cost or pay-off in cents for look:	1	+3	−1	+8	−2	+2	+9
		2	0	1	3	2	2	3
		3	2	1	4	3	3	3
		4	1	1	4	3	3	3
		5	2	1	4	3	3	5
		6	1	2	5	3	3	4
		7	3	2	6	4	4	5
		8	2	2	6	4	4	5
		9	4	7	2	5	5	7
		10	4	8	3	5	5	6
		11	6	10	3	6	6	9
		12	7	12	3	7	7	10
		13	11	16	5	9	9	15
		14	17	24	8	12	12	22
		15		48		18		

a hole in the plastic sheet and found that his pencil punched through the paper instead, it meant that he had found the unique cell. He had no way of recognizing the unique cell without writing in it. It was possible by pressing very lightly with the pencil to identify the unique cell tactually without making a conspicuous mark. Close supervision of *S*s to prevent this and other kinds of cheating was necessary.

All *S*s were seated around the outside of a *U*-shaped group of tables. In front of each was a change-box containing $2.25 in poker chips, an instruction-sheet, a pencil, and four punchboards. The instruction-sheet contained complete information about rules, costs, and payoffs for the particular session. After reading it, *S* worked

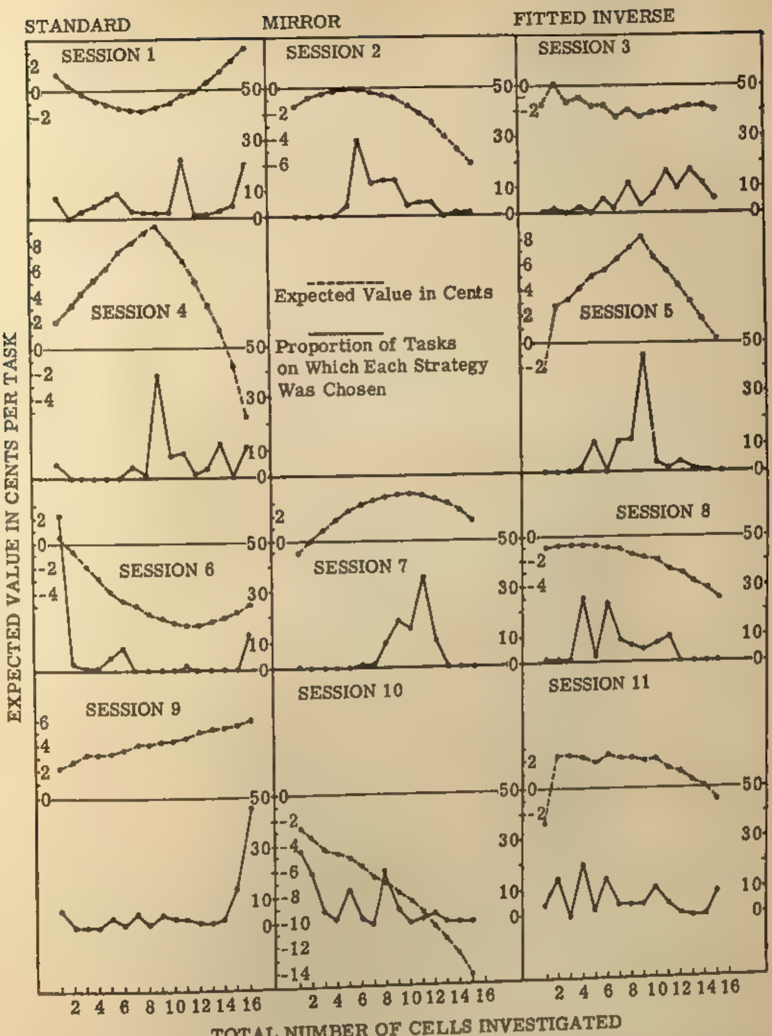

FIG. 1. EXPECTED VALUE OF EACH STRATEGY AND PROPORTION OF CHOICES OF EACH STRATEGY FOR EACH SESSION

The left-hand ordinate is the expected value in cents. The right-hand ordinate is the proportion of choices of each strategy.

through the 48 problems. One or two *E*s stood inside the *U*-shaped tables, monitoring the *S*s, and making and receiving payments. After the session, *S* returned \$0.75 in poker chips or cash, and any chips left he exchanged for money. In session 2, no chips were returned, and in sessions 10 and 11 all \$2.25 was returned; these minor variations were intended to even out the pay-off rate a little. Answer-sheets were checked at the end of each session to make sure that each *S* had reported his net win or loss accurately; thus there was no incentive to cheat by failing to make the payments specified by the rules of the game. After the experiment was completed, the *S*s if necessary attended additional sessions with cost- and pay-off-functions designed to bring their net winnings within the range of \$0.90-\$1.65 per hour. During the experiment itself, most *S*s averaged somewhat better than that.

Results

Strategy-scores. The most difficult procedural problem of the whole experiment was the selection of a satisfactory dependent variable indicating how many cells each *S* intended to look at on each task. That information was directly available for tasks in which *S* stopped looking without finding the unique cell, but not for those in which he found it. Unfortunately, the number of finds is directly correlated with *N*, *S*'s strategy for the task. In fact, if $N = 16$ for a standard task, a find is sure to occur. The solution was to extrapolate from trials in which finds did not occur to those in which finds did occur, assuming that within a single sequence of 48 tasks *S* made as few changes as possible, given the data, from one strategy to another. The rules for extrapolation are defined by the following algorithm:

(1) For each task in which the unique cell is not found, Strategy Score (*StS*) equals the total number of cells examined, including free or compulsory looks. For standard tasks in which the unique cell is found on the free guess, *StS* equals the total number of cells examined.

(2) For each task in which the unique cell is found, *StS* is equal to or greater than the number of cells examined. For standard tasks in which the unique cell is found as the result of a purchase rather than a free guess, *StS* is greater than the number of cells examined.

(3) Within the restrictions of Rules (1) and (2), *StS* are so chosen as to minimize the number of transitions from one *StS* to another within a sequence of 48 tasks all having the same costs and payoffs.

(4) Within the restrictions of Rules (1), (2), and (3), when a transition from one *StS* to another is necessary, it is placed as early as possible in the sequence of 48 tasks.

(5) Within the restrictions of Rules (1), (2), (3), and (4); the smallest permissible *StS* are chosen.

The relative importance of these rules varies with *StS*. For low *StS*, Rule (1) does most of the work. For high *StS*, Rules (2) and (3) are the most important. To illustrate the application of these rules, consider a sequence of inverse tasks.

The symbol n_s means that S stopped after n non-compulsory looks; the symbols n_f means that he found the unique cell on the nth non-compulsory look.

Response:	5_s	2_f	7_s	8_f	9_f	8_f	2_f	7_s	3_f
StS:	6	8	8	10	10	10	8	8	8
Key Rule:	1	3&4	1	3	5	3	4	1	3

In principle, Rule (3) could reduce seriously the variance of *StS* compared with that of *S*'s actual strategies, but the data give no reason to suppose that this actually happened. In particular, from task to task within a 48-task sequence, low *StS* vary no more than high ones, even though Rule (3) is seldom used for low *StS*. In fact, the correlation between mean strategy-scores over a 48-task sequence and the *SD*s of those means is 0.002, and the scatterplot is as scattered as that correlation suggests.

Quality of performance. Fig. 1 shows the distribution of *StS* for each of the 11 sessions, averaged over *S*s. The expected value of each strategy for each session is included for comparison. Inspection makes it obvious that the strategy chosen was in general closely related to its expected value. The optimal strategy was chosen more often than any other in 8 of the 11 sessions, and in the other three it was nearly the best-liked one. Close inspection of Fig. 1 also suggests that in standard-tasks deviations from the optimum were more likely to be in the direction of buying too much information, and that in inverse tasks they were more likely to be in the direction of looking at too few cells.

Fig. 2 highlights the over-all degree of optimality found by showing how many strategies were within specified distances from the optimal strategy; here and elsewhere in this paper, such distances are most appropriately measured in cents. The optimal strategy was chosen about 34% of the time. Over 45% of the *StS* are among the three best strategies.

Analysis of deviations from optimality. A deviation-score, defined as the difference in cents between the expected value of the chosen *StS* and the expected value of the optimal strategy, was calculated for each task. Table II shows the main effect of the sessions. The deviation-score which would have been expected by an *S* who chose at random from the strategies available in each task is included for comparison. The table shows again that the strategies chosen were better than chance, except in Session 3. This is true for each *S*. It shows also that the amount of obtained deviation is closely related to the amount that would have been obtained by chance, the correlation being 0.75. This finding is in part a consequence of differences in inter-strategy spread among the tasks. If the difference

between best and worst strategy is small, then both the maximal deviational score and the random-choice deviational score are inevitably small. Inspection of Fig. 1 makes it clear, however, that *Ss* also tended to stick closer to the optimal strategy when deviations from it were more costly than when they were less costly. It is worth noting that, in general, performance was better on inverse tasks (mean deviation 1.06¢) than on standard tasks (mean deviation 2.36¢).

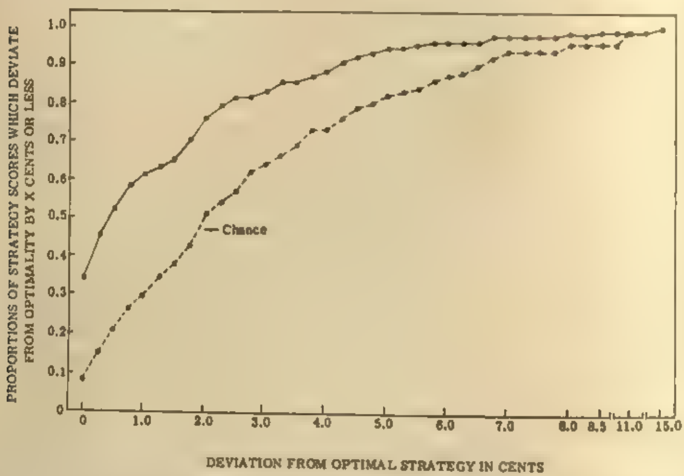

FIG. 2. CUMULATIVE DEVIATION FROM OPTIMAL STRATEGY AS A FUNCTION OF SIZE OF DEVIATION IN CENTS

The solid line represents the strategies which *Ss* actually used. The dashed-line labeled 'chance' represents the functions which would have been obtained if all available strategies has been used equally often.

Cautiousness. There are two ways of deviating from the optimal strategy. In standard tasks, an *S* who buys too many items of information before guessing is being too cautious, and one who buys too few is being too incautious. In inverse tasks, an *S* who looks at too few cells before stopping for fear of hitting the unique cell is being too cautious, while one who looks at too many is being too incautious. It is not possible for an *S* to be either too cautious or too incautious on all tasks. In order for both possibilities to exist, the optimal strategy must fall somewhere in the middle of the range of strategies. This is the case only for Sessions 2, 4, 5, and 7. For those four neutral sessions, a cautiousness-measure was defined as a ratio of deviation from the optimal strategy in cents attributable to excessive caution to total deviations from the optimal strategy for these sessions.

Fortunately, the *Ss* divide exactly into five cautious and five incautious *Ss*, whether the dividing line is the median, the 0.76 mean-cautiousness, the 0.57 mean of the cautiousness-ratios, the 0.52 random-choice strategy, or the 50% cautiousness-ratio. Of the seven sessions in which the optimal strategy does not fall near the middle of the range of available strategies, four (Sessions 6, 8, 9, and 10) favor high cautiousness and the other

TABLE II

MEAN DEVIATION FROM OPTIMALITY IN EACH SESSION

Session no.	Task type	Observed deviation (in cents) per task	Random-choice deviation (in arts)
9	standard	0.76	1.72
6	standard	1.86	5.00
1	standard	2.67	3.12
4	standard	4.15	5.23
Average		2.36	3.77
7	inverse	0.13	1.29
2	inverse	0.69	1.54
10	inverse	2.07	4.87
8	inverse	0.51	1.21
11	inverse	0.97	1.32
5	inverse	1.18	3.84
3	inverse	1.86	1.66
Average		1.06	2.25

TABLE III

DEVIATIONS FROM OPTIMAL STRATEGY IN CENTS PER TASK FOR CAUTIOUS AND INCAUTIOUS *Ss* ON NEUTRAL AND CAUTION-FAVORING TASKS

	Neutral tasks (Sessions 2, 4, 5, 7)	Caution-favoring tasks (Sessions 6, 8, 9, 10)
Cautious *Ss*	2.19	0.91
Incautious *Ss*	1.01	1.76
Random-choice strategy	2.98	3.21

three (Sessions 1, 3, and 11) are ambiguous or bimodal. Table III compares the deviations from optimality for cautious with those for incautious *Ss*, and the effect of the random-choice strategy included for comparison. The incautious *Ss* do better than the cautious *Ss* in the neutral sessions, and worse in the sessions which favor high caution; the random-choice strategy does equally badly in both types of session. This finding indicates that *Ss* have a style for performing these information-seeking tasks that can be described, at least to some extent, in terms of cautiousness or incautiousness, that is recognizable over changes in cost-payoff conditions, and that causes better performance when style and task match than when they conflict.

Discussion

By far the most important finding in this experiment is the surprisingly high quality of performance of a difficult task. We are willing to assume that no *S* was able to calculate the expected value of any strategy; as far as we could tell, none tried. Nevertheless, many choices were for optimal strategies, and most of the rest were for near-optimal ones. How could *S*s do so well? Inspection of Table I gives some clues. Consider, for example, Session 5, an early inverse session for which the incidence of optimal strategies was nearly 50%. The pay-off for each successful look rises from 3 to 6 cents, and then suddenly drops down to 2 cents at the ninth look. An *S* might or might not take as many as eight looks, but it seems very unlikely that a reasonable man, having taken eight successful looks, would take the ninth. A similar remark with reversed signs applies to Session 4, for which performance also was very good, but performance was excellent also in Sessions 2 and 7, in which there were no such simple cues. Session 6 and 9 seem to produce excellent performance by having extreme optimal strategies. They both are standard-task sessions, but Session 6 favors incaution and Session 9 favors caution. In Sessions 3 and 11, both inverse sessions with variable pay-offs per look, performance is rather poor, while in Session 8, also an inverse session but with fixed pay-offs per look, it is relatively good. This finding suggests that sessions with fixed costs or pay-offs per look are easier to understand than those in which the costs or pay-offs are functions of the number of looks, but this hypothesis cannot account for the excellent performance in Sessions 4 and 5, or for the relatively poor performance in Session 1. In short, no appealing simple explanation of the good performance, or of the deviations from it, seems to stand up under close examination. We therefore tentatively conclude that men simply do a good job of intuitively sizing up strategies in relatively complex tasks of information-seeking. The reason may be that they have had practice at such tasks, under aperiodic reinforcement, with rewards and punishments much more potent than those usable in the laboratory.

This conclusion is not new. Indeed, it is the most familiar conclusion in the relatively small literature of information-seeking.[2] It seems somewhat at variance with the findings of the many experiments on choices among bets,[3] but the disagreement is not real. Data on choices among

[2] See, for example, F. W. Irwin and W. A. S. Smith, Value, cost, and information as determiners of decision, *J. exp. Psychol.*, 54, 1957, 229-232; D. G. Pruitt, Informational requirements in making decisions, this Journal, 74, 1961, 433-439; and G. M. Becker, Sequential decision making: Wald's model and estimates of parameters, *J. exp. Psychol.*, 55, 1958, 628-636.

bets require notions like utility and subjective probability for their interpretation primarily because few if any such experiments ask the *S*s to choose among bets whose expected values differ significantly from one another. Experimenters do not do such experiments because they are confident of what will happen: *S* will prefer the bet with the highest expected value. The only reason the present experiment was worth doing was that this common-sense prediction seemed less self-evident for so complex and sequential a task. The data indicate that the prediction, though less self-evident, still holds. When given the opportunity, men most often maximize expected value.

When they do otherwise, they can be considered to have a choice between excessive caution (preferring high probabilities of small gains) and excessive incaution (preferring low probabilities of high gains). The data of this experiment indicate that an *S* who is cautious in one situation is likely to be cautious in another, similar one, and thus raise the question of how general a personality trait caution or incaution may be. If stable individual differences in information-seeking tasks can be demonstrated to exist, the high face-validity of this task, along with the ease of administering and scoring, suggests that it may be useful in selecting decision-makers.

Summary

Undergraduate *S*s (10 men) were asked to perform 528 information-seeking tasks of two types. In standard tasks, *S* paid for looks at the cells of a 16-cell matrix and was rewarded if he found the unique cell. In inverse tasks, *S* was fined for finding the unique cell and rewarded for each cell investigated which was not the unique one. The basic dependent variable, strategy-score, was the number of cells *S* looked at for tasks in which he did not find the unique cell. For tasks in which he did find it, elaborate rules were used to infer what he would have done had he not found it. The *S*s performed remarkably well; about half the strategies used were optimal or approximately so, and serious divergences from the optimum were very rare. Performance was less good on standard than on inverse tasks, and slightly less so on tasks with more complex costs and pay-offs than on tasks with simpler costs and pay-offs. Half the *S*s were too cautious and half were too incautious in tasks for which both were possible. Individual *S*s usually were consistent in being cautious or incautious.

---

[3] For reviews of such experiments, see Ward Edwards, The theory of decision making, *Psychol. Bull.*, 51, 1954, 380-417; Behavioral decision theory, *Annu. Rev. Psychol.*, 12, 1961, 473-498.

PARTIAL REINFORCEMENT IN THE FISH: EXPERIMENTS WITH SPACED TRIALS AND PARTIAL DELAY

By R. C. Gonzalez, Erika R. Behrend, and M. E. Bitterman, Bryn Mawr College

The results of the first experiments on partial reinforcement in the fish were quite different from those obtained in analogous experiments with the rat: Consistently reinforced animals were more resistant to extinction than partially reinforced animals given the same number of training trials.[1] In subsequent experiments, with number of reinforcements equated, resistance was greater after partial reinforcement, but the possibility remained that the effect was due to the greater number of training trials given the partial animals rather than to partial reinforcement *per se*.[2] To demonstrate unequivocally that resistance to extinction in the fish is increased by partial reinforcement required a further experiment, in which two partial groups had the same number of training trials and the same number of reinforcements but different schedules of reinforcement; a group trained with long runs of unreinforced trials showed greater resistance than did a group trained only with short runs of unreinforced trials.[3] It should be noted that the partial groups in the early equated-trials experiments had only short runs of unreinforced trials in training —too short, apparently, to outweigh the difference in frequency of reinforcement which favored the consistent groups. With longer runs, greater resistance after partial reinforcement may be expected in the fish, even when trials rather than reinforcements are equated.

---

* Received for publication September 4, 1963. This investigation was supported by Grants MH-02857 and MH-9577 from the National Institute of Mental Health.

[1] Jerome Wodinsky and M. E. Bitterman, Partial reinforcement in the fish, this Journal, 72, 1959, 184-199; Resistance to extinction in the fish after extensive training with partial reinforcement, this Journal, 73, 1960, 429-434; Nicholas Longo and M. E. Bitterman, The effect of partial reinforcement with spaced practice on resistance to extinction in the fish, *J. comp. physiol. Psychol.*, 53, 1960, 169-172.

[2] R. C. Gonzalez, R. M. Eskin, and M. E. Bitterman, Extinction in the fish after partial and consistent reinforcement with number of reinforcements equated, *ibid.*, 55, 1962, 381-386.

[3] Gonzalez, Eskin, and Bitterman, Further experiments on partial reinforcement in the fish, this Journal, 76, 1963, 366-375. A like effect has been found in the rat by Gonzalez and Bitterman, Resistance to extinction in the rat as a function of percentage and distribution of reinforcement, *J. comp. physiol. Psychol.*, 1964, 258-263.

Although it now seems clear that partial reinforcement tends to increase resistance to extinction in the fish as well as in the rat, there is some doubt that the underlying processes are the same. The doubt is based on divergent results obtained in two kinds of experiment which were designed originally to test the parsimonious Hull-Sheffield sensory-carryover interpretation of the effect of partial reinforcement on resistance to extinction in the rat.[4] Resistance to extinction in the rat is increased by partial reinforcement whether the intertrial interval is short or long—the interval may be as long, actually, as 24 hr.[5] In the fish, by contrast, no such increase was found in a one-trial-a-day experiment, even with equated reinforcements.[6] Resistance to extinction in the rat is greater after training with partial delay of reinforcement (which is like pure partial reinforcement except that long-delayed reinforcement is substituted for nonreinforcement) than after training with consistent reinforcement.[7] In the fish, by contrast, no such difference has been found.[8] The results for the rat are difficult to explain in terms of sensory carryover, but they do not mean that sensory carryover does not operate at all. They suggest only that there must be some additional mechanism (or mechanisms) operating in the rat to produce a like effect, while the results for the fish do not require us to assume any other mechanism. The experiments now to be reported were designed to provide further information on partial reinforcement in the fish under conditions which have given different results for the two animals.

---

[4] F. V. Sheffield, Extinction as a function of partial reinforcement and distribution of practice, *J. exp. Psychol.*, 39, 1949, 511-526; C. L. Hull, *A Behavior System*, 1952, 120-121.

[5] Solomon Weinstock, Resistance to extinction of a running response following partial reinforcement under widely spaced trials, *J. comp. physiol. Psychol.*, 47, 1954, 318-322; Wilma Wilson, E. J. Weiss, and Abram Amsel, Two tests of the Sheffield hypothesis concerning resistance to extinction, partial reinforcement, and distribution of practice, *J. exp. Psychol.*, 50, 1955, 51-60; D. W. Tyler, Extinction following partial reinforcement with control of stimulus-generalization and secondary reinforcement, this JOURNAL, 69, 1956, 359-368; D. J. Lewis, Acquisition, extinction, and spontaneous recovery as a function of percentage of reinforcement and intertrial intervals, *J. exp. Psychol.*, 51, 1956, 45-53.

[6] O. C. Shaffer, Resistance to extinction in the goldfish following partial reinforcement in an instrumental situation. Unpublished Master's thesis, Bryn Mawr College, 1961.

[7] Janet Crum, W. L. Brown, and M. E. Bitterman, The effect of partial and delayed reinforcement on resistance to extinction, this JOURNAL, 64, 1951, 228-237; E. D. Scott and E. L. Wike, The effect of partially delayed reinforcement and trial-distribution on the extinction of an instrumental response, this JOURNAL, 69, 1956, 264-268; F. A. Logan, E. M. Beier, and W. D. Kincaid, Extinction following partial and varied reinforcement, *J. exp. Psychol.*, 52, 1956, 65-70; Elizabeth Fehrer, Effects of amount of reinforcement and of pre- and post-reinforcement delays on learning and extinction, *J. exp. Psychol.*, 52, 1956, 167-176.

[8] Gonzalez, Eskin, and Bitterman, *op. cit.*, this JOURNAL, 76, 1963, 366-375.

Experiment I: Spaced Trials

Although all other instrumental experiments on partial reinforcement in the fish were done with the African mouthbreeder, the equated-reinforcements, spaced-trials experiment that failed to produce the partial-reinforcement effect was done with the goldfish. Since there is still some question about whether the goldfish will show the partial-reinforcement effect even in massed trials,[9] it is of interest to look at some parallel results for the mouthbreeder. In the experiment with the goldfish, only short runs of unreinforced trials were scheduled in training. In the experiment with the mouthbreeder now to be reported, two partial groups were used, one given relatively long runs of unreinforced training trials, and a second given only short runs.

Subjects. The *S*s were 36 experimentally näive mouthbreeders, about 3 in. long, which had been bred in the laboratory. They lived in individual 2-gal. tanks set on open shelves in a temperature-controlled room.

Apparatus. The apparatus employed has been described in detail elsewhere.[10] It consisted of a target which could be lowered into *S*'s living tank, an electronic system for detecting *S*'s contacts with the target, and a pellet-feeder operated by the detecting system. A Standard Electric timer, graduated in units of 0.01-sec., was used to measure latency of response; the clock started with the introduction of the target and was stopped by *S*'s response.

Preliminary training. After adaptation to the experimental situation, the *S*s were trained to strike at a target baited with Aronson's mixture. Each *S* was brought in its living tank to the experimental situation. After 5 sec., the baited target was introduced, and response to it was rewarded with 15 pellets of food (the amount of reward given throughout the experiment). Immediately after response, the target was withdrawn and, when *S* had consumed the food, it was returned to its living area. This procedure was continued daily until *S* was responding readily (which required about eight days on the average), and then training with an unbaited target was begun. There were 10 reinforced trials with the unbaited target, one per day, on each of which the latency of response was recorded. On the basis of their performance during these 10 days, the animals were divided into three matched groups of 12 *S*s each: A Consistent Group, a Gellermann Group, and an Extended Group.

Experimental training. The first stage of the experiment proper lasted 50 days. Again, each *S* had one trial per day. The Gellermann and the Extended animals were given 50 trials in all, 25 reinforced and 25 unreinforced. For the Gellermann Group, the two kinds of trial were always scheduled according to Gellermann's rules (which limit the number of reinforced or unreinforced trials in succession

---

[9] Barry Berger, Matthew Yarczower, and M. E. Bitterman, The effect of partial reinforcement on the extinction of a classically conditioned response in the goldfish, *J. comp. physiol. Psychol.*, in press.

[10] Longo and Bitterman, Improved apparatus for the study of learning in fish, this Journal, 72, 1959, 616-620.

to three). For the Extended Group, Gellermann's rules were used on the first 10 days only; thereafter, the order was changed to involve progressively longer runs of unreinforced trials up to a maximum of eight (on Days 38-45). The exact sequence, with *R* standing for reinforcement and *N* for nonreinforcement was: NRNRR NNNRR RNNNN RRRRR NRRRN NNNNN RRRNR RRNNN NNNNN RRRRR. The Consistent Group, which had 25 reinforced trials but no unreinforced trials, was divided into two matched subgroups, one of which had trials on days when the Gellermann Group was reinforced, and the other on days when the Extended Group was reinforced. Throughout this as well as all other stages of the experiment, the animals were maintained on a feeding schedule of 20 pellets per day. That portion of the daily ration which was not given in the experimental situation was given in the living room area 1 hr. after each day's trial.

In the second stage of the experiment, each *S* was given one unreinforced trial per day. When, on any trial, an animal failed to respond in 60 sec., the target was withdrawn, the time was recorded, and *S* was returned to its living area. Work with each animal was terminated after five successive failures to respond in 60 sec.

Results. The two Consistent subgroups were combined for purposes of analysis because they did not differ from each other in acquisition or in extinction. All analyses are based on the data for 12 Consistent, 11 Gellermann, and 10 Extended *S*s, 3 animals having been lost for one reason or another in the course of the experiment.

The latency of response was already very low (about 0.1 sec.) at the end of the pretraining, and it declined very little during the first stage of the experiment. The differences among the three groups in this stage were negligible ($F < 1$). The long runs of unreinforced trials given the Extended Group had no measurable effect.

The performance of the three groups during extinction is plotted in Fig. 1 in terms of the mean number of trials to successive criteria of extinction (one, two, three, four, and five consecutive failures to respond in 60 sec.). There is no suggestion in these curves of a tendency for partial reinforcement to increase resistance to extinction. The curve of the Gellermann Group is lower than that of the Consistent Group, while the curve of the Extended Group is the lowest of all. A repeated-measures analysis of variance indicates a significant Groups-effect ($F = 3.4$, $df = 2/30$, $p < 0.05$), which is accounted for entirely by the low resistance of the Extended Group (for the difference between the Extended Group and the two other groups combined, $F = 6.69$, $df = 1/30$, $p < 0.05$; for the difference between the Gellermann Group and the Consistent Group, $F < 1$). In Fig. 2, the course of extinction to the first criterion is shown in terms of median log latency (plus 2) for blocks of five trials, each curve being plotted to the point at which the median animal of the group failed for the first time to respond in 60 sec. The curve for the Extended

Group diverges from those of the other two groups quite early in extinction, beginning after the fifteenth trial, while the Gellermann and the Consistent curves overlap throughout.

Discussion. These results differ from those of massed-trials, equated-reinforcements experiments with the mouthbreeder in that partial rein-

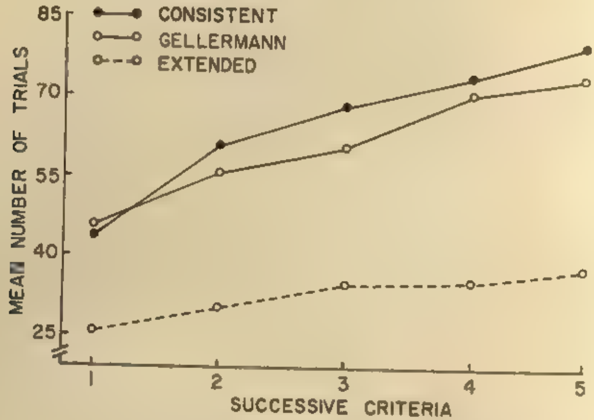

Fig. 1. Trials to Successively More Severe Criteria of Extinction in Experiment I

forcement failed to increase resistance to extinction. The difference may not, of course, be due to spacing, but to some confounded factor, such as, for example, the relatively small number of training trials. It should be noted, however, that the partial-reinforcement effect has been found in a parallel experiment with the pigeon,[11] as well as in a one-trial-a-day experiment with the rat in which there were not many more trials and even fewer reinforcements.[12]

The present results also differ from those of massed-trials experiments with the mouthbreeder in that long runs of unreinforced training trials produced less resistance to extinction than did short runs, and here again number of training trials rather than spacing may have been responsible. Long runs of unreinforced trials tend to produce extinction; if resistance ultimately is to be increased by long runs, it is reasonable to think that the animal must receive a good deal of reinforcement for response under

[11] W. A. Roberts, D. H. Bullock, and M. E. Bitterman, Resistance to extinction in the pigeon after partially reinforced instrumental training under discrete-trials conditions, this Journal, 76, 1963, 353-365.

[12] Weinstock, *op. cit.*, 318-322.

conditions of partial extinction. In the parallel experiment with the pigeon already cited, resistance to extinction was no greater in the Extended Group than in the Gellermann Group, although both showed greater resistance than did the Consistent Group. It is possible, however, that long runs are especially effective in massed trials because the aftereffects of

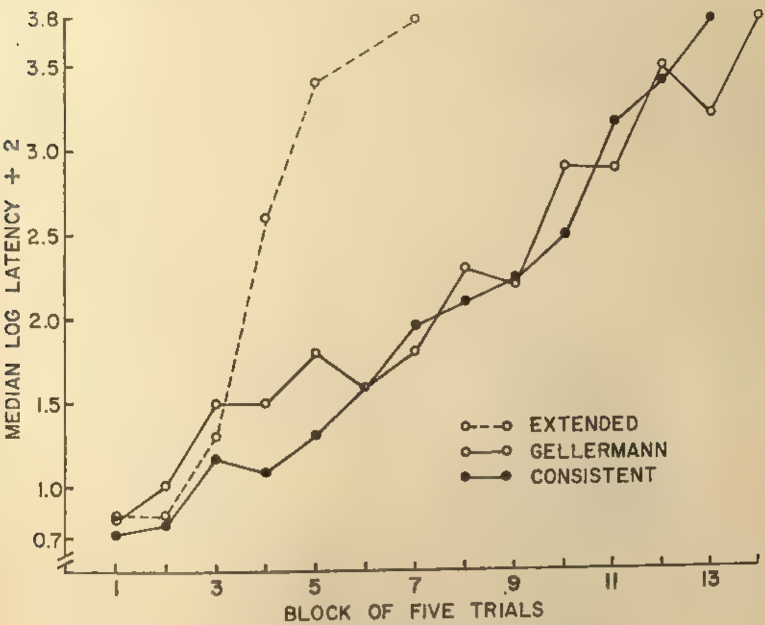

FIG. 2. THE EARLY COURSE OF EXTINCTION IN EXPERIMENT I

nonreinforcement cumulate to produce a set of sensory conditions very much like those to be encountered later in extinction. As matters now stand, in any case, there is nothing in the data for the fish which calls for a mechanism more complex than sensory carryover.

EXPERIMENT II: PARTIAL DELAY

In the one previous experiment on partial delay of reinforcement in the fish, only short runs of delay-trials were scheduled in training—Gellermann orders were used; but the same orders produced only a rather weak effect under conditions of pure partial reinforcement.[13] In the experiment now to be reported, longer runs of delay-trials were scheduled.

[13] Gonzalez, Eskin, and Bitterman, *op. cit.*, this JOURNAL, 76, 1963, 366-375.

Subjects. The *Ss* were 30 experimentally näive mouthbreeders, about 3 in. long, bred in the laboratory and maintained under the same conditions as those of the previous experiment.

Apparatus. The apparatus was the same as that used in Experiment I.

Preliminary training. After adaptation to the experimental situation, the *Ss* were trained to strike at a continuously available target baited with Aronson's mixture. Each contact of *S* with the target was rewarded with a pellet of food until 10 pellets had been earned. This procedure was repeated daily until *S* was responding readily, which required three days on the average, and then training with an unbaited target was begun. There were six daily sessions with the unbaited target, each consisting of 20 reinforced, discrete trials; the target was withdrawn after each response and introduced again after reinforcement. On the basis of their performance during these six days, the *Ss* wre divided into two matched groups of 15 each—an immediate (Consistent) Group and a Partial-Delay Group.

Experimental training. The training-stage of the experiment consisted of 30 daily sessions of 20 trials each. For the Immediate Group, each response was reinforced immediately. For the Partial-Delay Group, half the responses were reinforced immediately and the remainder were reinforced after a delay of 30 sec. (On delay-trials, the target was withdrawn from the water immediately after response, and an interval timer, which was started by the response, activated the pellet-feeder after completing its cycle.) On the first day of training, the schedule of immediate and delay-trials followed Gellermann's rules; thereafter, the schedule was changed to involve progressively longer runs of delay-trials, up to a maximum of 11 in a single day and 15 over a block of two days. The longest run of delay-trials was 2 on Day 1, 5 on Day 2, and 7, 9, 9, 9, 4, 9, 6, 7, 8, 10, 6, 9, 4, 11, 10, 5, 6, 9, 5, 10, 8, 7, 6, 4, 11, 6, 9, and 9 on Days 3-30, respectively. The point within each day's block of 20 trials at which the long runs began was varied irregularly; sometimes it was early in the session, sometimes in the middle, and sometimes toward the end. At 10 points in training, a run scheduled late in a session was extended into the early trials of the following day (*e.g.,* Days 8 and 9), making the length of run over each of these two-day blocks 15. The trials were massed for both groups, the interval between reinforcement and subsequent introduction of the target averaging about 3 sec.

On Day 31, all *Ss* were extinguished. On each extinction-trial, the target was withdrawn from the water immediately after response, and then reinserted 3 sec. later to begin the next trial. When, on any trial, *S* failed to respond in 30 sec., the target was withdrawn, the time was recorded, and the target was reinserted in the usual manner 3 sec. later. The criterion of extinction was five successive failures to respond in 30 sec.

Results. Latency of response in acquisition declined in negatively accelerated fashion for both groups to stable levels—about 0.16 sec. in the Immediate Group and 0.30 sec. in the Partial-Delay Group—which were reached by Day 15. Analysis of variance, based on the sum of the individual scores for the last 15 days of training, showed a significant effect of partial delay of reinforcement on asymptotic latency ($F = 4.32$,

$df = 1/28$, $p < 0.05$). Early in training, the effect was due almost entirely to the tendency for long runs of delayed reinforcements to produce progressive increases in latency of the sort which characterize extinction. As training continued, however, this tendency diminished somewhat, although the over-all latency of the Partial-Delay Group remained relatively high.

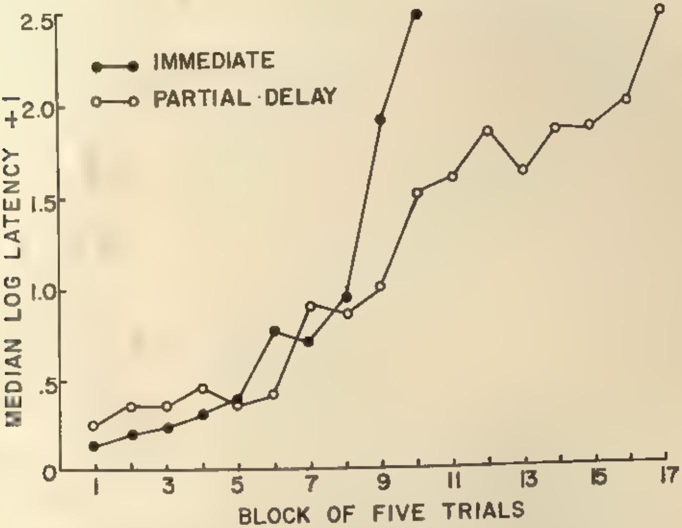

FIG. 3. THE EARLY COURSE OF EXTINCTION IN EXPERIMENT II

In Fig. 3, the course of extinction to the point at which the median *S* of each group failed for the first time to respond in 30 sec. is plotted in terms of median log latency (plus 1) for blocks of five trials. The two curves overlap until the ninth block, when the first indication of greater resistance in the Partial-Delay Group begins to appear. A repeated-measures analysis of variance for the first 40 trials of extinction (the trials before the separation of the curves) yields a significant Trials-effect ($F = 7.24$, $df = 39/1092$, $p < 0.001$), indicating that the divergence of the curves begins only after both groups already have undergone a significant amount of extinction.

The entire course of extinction for the two groups is plotted in Fig. 4 in terms of mean number of trials to successively more severe criteria. Although the resistance of the Partial-Delay Group appears to be greater than that of the Immediate Group at the first criterion, the difference between the groups is not significant ($F = 2.37$, $df = 1/28$, $p > .05$). In

the further course of extinction, however, the difference in resistance becomes substantial. A two-way analysis of variance yields a significant Groups-effect ($F = 4.52$, $df = 1/28$, $p < 0.05$); a significant Criterion-effect ($F = 23.53$, $df = 4/112$, $p < 0.001$), and a significant interaction of Groups with Criteria ($F = 2.91$, $df = 4/112$, $p < 0.05$).

Discussion. The principal importance of this experiment is that it provides a demonstration of greater resistance to extinction in the fish after

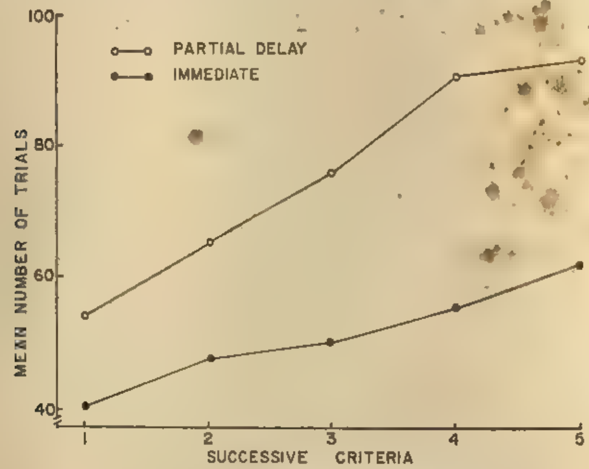

FIG. 4. TRIALS TO SUCCESSIVELY MORE SEVERE CRITERIA OF EXTINCTION IN EXPERIMENT II

inconsistent as compared with consistent reinforcement under *equated-trials* conditions. Just how the effect should be interpreted is, however, in doubt. If in delayed reinforcement all after-effects of nonreinforcement are replaced by after-effects of reinforcement, then the greater resistance of the Partial-Delay Group cannot be accounted for in terms of sensory carryover, and some additional mechanism must be assumed for the fish as well as for the rat. Anticipating the possibility that the partial-reinforcement effect might occur in spaced trials, Sheffield suggested that the sensory consequences of nonreinforcement might be *associatively reinstated* after a long intertrial interval.[14] Her notion, which applies as well to partial delay, was that the frustration produced by nonreinforcement might be conditioned to apparatus cues and evoked again by those cues

[14] Sheffield, *op. cit.*, 521.

at the beginning of each trial. As an alternative to conditioned frustration, the somewhat broader notion of *extinctive carryover* has been suggested.[15] If nonreinforcement or long-delayed reinforcement produces some extinction, and if the extinction itself gives rise to certain characteristic sensory elements, then reinforcing the responses of a partially extinguished fish in training will serve to encourage response under the conditions of partial extinction which will be met again some time after all reinforcement has been terminated. Either notion will account for the present results.

It must be recognized, however, that the partial-delay experiment does not provide as powerful a test of the sensory-carryover interpretation as does the spaced-trials experiment. While it seems reasonable to assume that the sensory consequences of nonreinforcement dissipate entirely over a 24-hr. interval, there is room for doubt that they are obliterated completely by a delayed reinforcement. If some of the frustration generated between response and reinforcement carries over to the next trial in spite of the interpolated reinforcement and cumulates over long runs of delay-trials, the resulting increase in resistance to extinction can be explained in terms of the Hull-Sheffield hypothesis. From this point of view, there still is no need to go beyond sensory carryover at the level of the fish.

Summary

Two experiments on partial reinforcement in the mouthbreeder are reported. In the first experiment, three groups of *S*s were trained with equated-reinforcements and a 24-hr. intertrial interval—a Consistent Group and two partial groups, Gellermann and Extended. The Gellermann Group, which had only short runs of unreinforced trials in training, was no more resistant to extinction than was the Consistent Group, and the Extended Group, which had long runs of unreinforced trials in training, was significantly less resistant than the other two. In the second experiment, with massed trials, a Partial-Delay Group given long runs of training trials on which reinforcement was delayed showed significantly greater resistance to extinction than did a consistently reinforced Immediate Group given the same number of training trials. Both results are compatible with the Hull-Sheffield sensory-carryover interpretation.

[15] Gonzalez, Eskin, and Bitterman, *op. cit.*, this JOURNAL, 76, 1963, 374; Gonzalez and Bitterman, *op. cit.*, 258-263.

BOREDOM-INDUCED CHANGES IN PREFERENCES AMONG BETS

PAUL SLOVIC, SARAH LICHTENSTEIN, and WARD EDWARDS,
University of Michigan

This paper examines a methodological dilemma of research on preferences among bets. As a by-product, it validates a single-stimulus method for studying such preferences called the bidding method, which is far more efficient than the customary method of paired comparisons. The dilemma is this: *S*s who are highly motivated to make careful choices use complex, frequently shifting strategies and so produce data which appear inconsistent, variable, and disorderly. Bored, unmotivated *S*s adopt simple, easy-to-discover strategies, and change them seldom. Those *S*s who play a high proportion of their choices, who are run individually, and who are not required to make too many choices per session tend to be highly motivated; *S*s required to make rapid, imaginary choices in large group, experimental sessions tend to be bored.

The dilemma and the effects of the two ways of resolving it are illustrated by comparing the imaginary group choice experiment of Coombs and Pruitt with the real-gambling experiments of Edwards.[1] Comparison of procedures and of findings concerning consistency, transitivity, choice of independent variables, and single-peakedness specifies the setting and hypotheses underlying this paper.

Procedure. Coombs and Pruitt ran all their *S*s in one large group of 99 *S*s. All choices were imaginary, in the sense that nothing depended on them, and the *S*s knew it. Each *S* had a booklet. on each page of the booklet two bets were written, and *S* simply checked the preferred bet. The *S*s made their choices at rates ranging from 3.4 to 22.4 choices per minute, and so cannot have taken much time for reflection. Edwards ran *S*s in most of his experiments one at a time, played each

* Received for publication August 23, 1962. The work here reported was sponsored by the United States Air Force under Contract AF 49(638)-769. Document Number AFOSR-1134.

[1] C. H. Coombs and D. G. Pruitt, Components of risk in decision making: Probability and variance preferences, *J. exp. Psychol.*, 60, 1960, 265-277; Some characteristics of choice behavior in risky situations, *Ann. N. Y. Acad. Sci.*, 89, 1961, 784-794. Ward Edwards, Probability-preferences in gambling, this JOURNAL, 66, 1953, 349-364; Probability-preferences among bets with differing expected values, *ibid.*, 67, 1954, 56-67; The reliability of probability-preferences, *ibid.*, 67, 1954, 68-95; Variance preferences in gambling, *ibid.*, 67, 1954, 441-452; The prediction of decision among bets, *J. exp. Psychol.*, 50, 1955, 201-214.

chosen bet immediately after the choice, and paid after each win or collected after each loss. His Ss typically made about 50 choices per hour.

Consistency. The Ss often fail to choose between the members of a pair of bets the same way on several repetitions of the choice; such failures are instances of inconsistency. Coombs and Pruitt repeated each choice eight times, and reported information measures which summarize the extent of consistency they found. Edwards typically repeated each choice four times. He obtained a substantial incidence of inconsistent choices, but reported no quantitative information about inconsistency.

Transitivity. Coombs and Pruitt focused their discussion of transitivity on two concepts, weak stochastic transitivity (*WST*) and strong stochastic transitivity (*SST*). If *SST* applies, then if $p(APB) \geq 0.5$ (read: the probability that bet A is preferred to bet B is at least 0.5) and $p(BPC) \geq 0.5$ it follows that $p(APC) \geq \max\ [p(APB),\ p(BPC)]$; the last term says that the probability that A is preferred to C must be at least as large as the larger of the other two probabilities.

Almost all triples of choices in the Coombs-Pruitt experiment satisfied *WST*; those which did not were excluded from subsequent analyses. Coombs and Pruitt interpreted the relatively few violations as by-products of inconsistency. About 25% of all their triples failed to satisfy *SST*. Edwards counted occurrences of intransitive triads of choices in which A is preferred to B, B to C, and C to A separately in each replication of all paired comparisons for each set of bets. He found about 17% of the maximal number of intransitive triples that a diabolic S, trying to maximize that number, could have produced. No direct comparison between these findings concerning intransitivity is possible because of the differences in procedures for data-analyses and in number of replications of each choice. Both sets of findings lead to the same over-all impression: the Ss are moderately transitive. Edwards's kind of analysis tends to produce an incorrect casual impression of less transitivity than Coombs and Pruitt's kind of analysis, especially if *WST* rather than *SST* is used as the basis for drawing conclusions about transitivity.

Choice of independent variable. A bet may appropriately be considered a distribution of possible amounts of money which can be won or lost. Like any other distribution, it can be represented by a series of quantities (pay-offs), each with a probability of occurrence. Like any other distribution, a bet has moments, and familiar statistical quantities may be defined as a function of these moments. Thus a bet has a mean or expected value; for a two-outcome bet, $EV = p\mathrm{A} + q\mathrm{B}$, where p and $q = 1 - p$ are the probabilities of the two outcomes and A and B are their pay-offs. Similarly, a bet has a variance ($\sigma^2 = pq(\mathrm{A} - \mathrm{B})^2$), a skewness ($Sk = (1 - 2p)/(pq)^{1/2}$, and a kurtosis ($K = 1/pq - 3$). Note that for two-outcome bets, skewness and kurtosis are functions only of probabilities, not of pay-offs. Any three of the quantities p, A, B, EV, σ^2, Sk, K completely specify a two-outcome bet, except that specification of any one of p, Sk, and K completely specifies the other two, and so at least two of the three parameters must be chosen from A, B, EV, and σ^2. Some combinations are impossible; for instance, if A and B are negative, EV cannot be positive. Thus a wide variety of formally equivalent specifications of the independent variables in two-outcome bets is possible; much argument in static decision theory has focused on selection of the 'best' set of independent variables. Edwards used p, EV, and B; Coombs and Pruitt preferred EV, σ^2, and Sk.

Single-peakedness. A particularly simple set of choice strategies is the single-peaked strategies in which one bet is best liked and others are progressively less liked as they get farther and farther away from the best-liked one along the main dimension on which bets are conceived to vary—in the experiments at hand, p, σ^2, or both. The simplest possible choice-strategies are the subset of this single-peaked set in which S ranks the bets consistently from one end of that dimension to the other, liking the extreme bet at one end best and the extreme bet at the other end the least. Coombs and Pruitt found that 59% of preference orderings for their bets were single-peaked; 43% of their orderings were simple ascending or descending ones. Edwards made no counts which are directly comparable, but inspection of his group preference functions makes it clear that for positive and zero *EV* bets the preference function almost always has at least two peaks and often three or four, and inspection of unpublished data for individual *S*s confirms this impression. Edwards used sets of eight bets, while Coombs and Pruitt used sets of either five or six bets. This implies that Edwards had a greater opportunity to discover multiply-peaked orderings than did Coombs and Pruitt, but the difference is far too large to be attributable to this fact alone. Coombs and Pruitt are interested in single-peakedness because it is a consequence of their theory about preferential choice. We are interested in it because simple single-peaked strategies and simpler ascending or descending order strategies can be applied mechanically, without thought; their simplicity should make these strategies especially attractive to bored, unmotivated *S*s.

Experiment I

An attempt was made to replicate the Coombs-Pruitt findings, using procedures still more boring than theirs. Ten sets of five bets each were selected; all bets had zero *EV*. Five of the sets held p constant and varied σ^2 by varying A and B; the other five held σ^2 constant and varied p. Bets in each set were combined in all of the five possible groups of four bets. Each group of four bets was written on a sheet of a 50-page booklet. Eighty-two college students were asked to rank-order the four bets on each page; each S rank-ordered each group of four bets three times.

Results. The results of this experiment are generally comparable with the results of the Coombs-Pruitt experiment, and so no detailed discussion will be presented. Two substantial differences between these and the Coombs-Pruitt results were, however, observed. These *S*s found their ranking task so boring and so silly that they loudly and frequently complained about its monotony; Coombs and Pruitt do not report such complaints The most important difference comes in incidence of single-peaked orderings; 70% of all our orderings were single-peaked, and 55% were simple ascending or descending orderings. The comparable figures for the Coombs-Pruitt experiment are 59% and 43%. These findings are consistent with the hypothesis that the main effect of boredom is to increase incidence of single-peaked preference orderings; the conditions of this experiment should have been and evidently were even more boring than those of the Coombs-Pruitt experiment.

Experiment II

A more stringent test of the hypothesis underlying this study requires comparison of conditions which should produce several levels of motivation using the same *S*s and bets.

Since paired-comparison experiments are relatively inefficient, in the sense that they require many observations to determine scale-positions for relatively few stimuli, we also wanted to examine a far more efficient single stimulus-method called the bidding method. Such multi-method experiments on the same *S*s always raise questions about order effects. Since Coombs and Pruitt have already studied the effect of putting their boring conditions first, and since Edwards experiments have indicated that order of application of various motivational conditions governing choices among bets makes little difference compared with the nature of the conditions themselves, we chose to use a single order for all *S*s, one which put the Coombs-Pruitt conditions last in the sequence of sessions.[3] Study of order effects as an independent variable would have required greatly increased amounts of experimentation with no reason to expect appropriate increases in yield of useful information. Randomization of order is not, within reasonable sizes for an experiment, a satisfactory way to eliminate order effects; it only makes it more difficult to make sense out of any order effects which may be obtained.

TABLE I

Design of Experiment II

Session	Method	Pay off condition (Fraction of bets played)	No. replications	Number stimuli presented
1	paired-comparisons	none	3	90 pairs
2	paired-comparisons	⅓	3	90 pairs
3	paired-comparisons	all	3	90 pairs
4	bidding method	none	4	90 bets
5	bidding method	⅓	4	90 bets
6	paired-comparisons	none	8	240 pairs

Method: (1) Subjects. The *S*s were seven volunteer undergraduate men.

Schedule. Each *S* participated in six experimental gambling sessions spaced over a two week interval. Table I shows the methods of collecting the data and the proportion of bets actually made during each session.

(2) Stimuli. The stimuli are shown in Table II. They consisted of five two-outcome bets at each of three *EV*-levels; positive, zero, and negative. All bets had a σ^2 of 4.0. The bets were stated in terms of the roll of a six-sided die. For example, a bet with $p = 2/6$ and $q = 4/6$ read "If you roll a 4 or a 5 you win \$2.82. If you roll a 1, 2, 3, or 6 you lose \$1.41." For the paired comparisons sessions, the *S*s chose one member of each of the 10 possible pairs of bets at each *EV*-level.

(3) The bidding method. For the single stimulus-bidding method of Sessions 4 and 5 each *S* was asked to state the largest amount of money he would be willing

[3] Especially Edwards, *op. cit.*, this Journal, 66, 1953, 349-364; *op. cit.*, *ibid.*, 67, 1954, 68-95.

to pay *E* in order to play each bet. For an undesirable bet, *S* stated the smallest amount *E* had to pay him before *S* would play the bet. *S* was told that this amount of money, his bid for a bet, would be compared with bids that had been collected from other *S*s in former experiments, and that his bid would be accepted and played only if it fell in the upper quarter of previous bids for the bet. This competitive bidding was a fiction used to motivate *S* to give careful responses; there was no such comparitive group. *E* actually accepted a random third of *S*'s bids, hence *S* did not receive any consistent feedback useful to him for developing a strategy of play. *E* never, however, accepted a bid lower than 50¢ below the *EV* of the bet; if a bid scheduled for acceptance was too low, the next acceptable bid thereafter was accepted instead. Session 4, the first bidding session, was for practice; and no bets were made until Session 5.

TABLE II

BETS USED IN EXPERIMENT II

Upper value in each cell is the amount which can be won; probability of winning it is shown at the top of column. Lower value in each cell is amount which can be lost; probability of losing equals $1-p$.

Expected value	Probability of winning (p) 1/6	2/6	3/6	4/6	5/6
0¢	$4.50 −$ .90	$2.82 −$1.41	$2.00 −$2.00	$1.41 −$2.82	$ .90 −$4.50
+75¢	$5.25 −$ .15	$3.57 −$ .66	$2.75 −$1.25	$2.16 −$2.07	$1.67 −$3.75
−75¢	$3.75 −$1.65	$2.07 −$2.16	$1.25 −$2.75	$ .66 −$3.57	$ .15 −$5.25

(4) Group-session. The *S*s were run individually in Sessions 1 through 5. The bets or pairs of bets were typed on cards and *S* announced his bid or choice. Session 6, however, was a group-session in which each *S* was given a 12-page booklet with 20 pairs of bets on each page, and indicated his preferences by checking the preferred member of each pair. In this group session, no bets were played, and *S* left as soon as he completed the booklet. These conditions closely resemble those of the Coombs-Pruitt experiment. One of the seven *S*s failed to attend Session 6.

(5) Pay-off. Each *S* was paid by the hour for the non-gambling sessions. In the gambling sessions, poker chips were exchanged during the playing of the bets, and at the session's end *S* received cash if he had won money or paid his losses out of his own pocket. If necessary, additional sessions for each *S* were scheduled after all data had been collected. The sets of bets used in these additional sessions were designed to ensure that over the experiment as a whole, no *S* won less than 90¢ per hour nor more than $1.65 per hour.

Results: (1) Consistency and transitivity. To obtain a measure of consistency, we computed the proportion of time *S* made the same choice each of the three times he encountered a particular pair in the course of the three replications; the entries in Table III are averaged over pairs and *S*s.

For the bidding sessions, consistency was inferred from hypothetical paired-comparisons derived by ranking the bids.

To obtain a measure of intransitivity, we counted the number of sets of three bets which, for each *S*, had the property that *A* is preferred to *B*, *B* is preferred to *C*, and *C* is preferred to *A*. Table III contains the results of these analyses. The main finding exhibited by Table III is that the *Ss* were somewhat more consistent and slightly more transitive in Session 6 than in earlier Sessions.

The findings concerning consistency cannot be compared with either Coombs and Pruitt's or Edwards's previous previous findings. The findings

TABLE III

Analysis of Consistency and Transitivity
(Based on data from 6 *Ss*.)

	Session					
	1	2	3	4	5	6
Consistency (3 replications)	.59	.59	.65	.55*	.55*	.71†
Intransitivity (every replication)	.18	.16	.17	—‡	—‡	.14

* Proportion based on the first three replications in the session.
† Proportion calculated by averaging the proportions from the first three and last three replications.
‡ Method imposes transitivity.

concerning transitivity are directly comparable with Edwards's findings; the 17% incidence of intransitivity under real gambling conditions is identical with the 17% incidence of intransitivity under those conditions in Edwards's first experiment.[3] Session 6 shows more transitivity than Session 1; our best guess on the basis of this and other evidence is that both boredom and practice enhance transitivity.

(2) Single-peakedness. We counted the incidence of single-peaked strategies, using a procedure identical with Coombs and Pruitt's. Table IV presents the results for the six sessions of Experiment II; comparable results from Experiment I and from the Coombs-Pruitt experiment are included for comparison.

Note the large proportion of the simple kinds of orderings generated in Session 6. While this increase indicates real changes, of preference, close examination of the data reveals that the *Ss* made the smallest possible alterations that would result in single-peaked orderings. For example, the

[3] Edwards, *op. cit.*, this Journal, 66, 1953, 349-364.

double-peaked order 5/6, 4/6, 3/6, 1/6, 2/6 found in Sessions 1 through 5 might change to 5/6, 4/6, 3/6, 2/6, 1/6 in Session 6.

To determine whether these proportions of simple orderings were a function of the number of replications, we made additional counts of the nine replications of Sessions 1, 2, and 3 combined and of the first and last three replications of Session 6. The proportions of simple orderings were not affected by number of replications.

(3) Preference-patterns. Vote-counts were obtained from the data of the paired-comparisons. A 'vote-count' is a count of the number of times each bet was chosen over all other bets with which it was paired. We

TABLE IV

ORDERING BY PROBABILITY-PREFERENCE

	No. *S*s	No. orderings	Orderings: Single peaked No.	Single peaked Propor.	Ascending or descending No.	Ascending or descending Propor.
Experiment II						
Sessions 1, 2, and 3	7	63	7	.11	4	.06
Sessions 4 and 5	7	42	3	.07	2	.05
Session 6	6	18	9	.50	5	.28
Coombs-Pruitt Study	99	198	116	.59	85	.43
Experiment I	82	410	287	.70	225	.55

made a comparable count for the bidding sessions by ranking bids and generating a hypothetical set of paired comparisons data from the rankings.

Group probability-preferences were examined on the basis of the vote-counts. For positive and zero *EV*-levels, the preference-pattern has peaks at 1/6 and 4/6; the least well-liked bet was the 2/6 bet. The peak at 1/2, found by Edwards and by Coombs and Pruitt, did not occur in these data; we don't know why not. The preference pattern for negative *EV*-bets looks rather like that for zero *EV*-bets. Edwards found strong preferences for low probabilities of loss and large amounts of loss for negative *EV*-bets; the divergence between those findings and these probably results from the fact that these negative *EV*-bets include a positive as well as a negative outcome, while the best possible outcome in each of Edwards's negative *EV*-bets was zero. Thus these negative *EV*-bets resemble zero *EV*-bets more closely than did Edwards's.

Group preference patterns for Session 6 increase monotonically from 1/6 to 5/6 for the zero *EV*-bets, and nearly do so for positive *EV*- and negative *EV*-bets also.

The bidding sessions produced preferences very similar to those ob-

tained in the first three paired-comparison sessions, except for one seriously divergent point (too low) for the 3/6 negative *EV*-bet. There were no differences worth noting between Sessions 2 and 3 or between Sessions 4 and 5.

(4) Bid-analysis. While the *S*s bid less for bets in Session 5, the actual gambling session, than in Session 4, the practice-session, the relative preferences were the same. In the actual gambling session, the *S*s bid an average of 33¢ for positive *EV*-bets, minus 23¢ for zero *EV*-bets, and minus $1.07 for negative *EV*-bets.

Bids can be analyzed simply by looking at mean bids, or by inferring a rank-order from the bids and then generating hypothetical paired comparisons from the rankings. No differences worth mentioning exist between the results of these procedures for the positive and zero *EV*-bets. For the negative *EV*-bets, a few *S*s defensively underbid for the 5/6 bet, which has a large possible loss, and a few *S*s bid more than usual, though still less than the *EV* of the bet, for the 1/6 bet, which has a large possible win and a comparatively low possible loss. Thus the mean bids were somewhat higher for the 1/6 bet and lower for the 5/6 bet than the vote count generated from the bids would lead you to expect.

Discussion. The first five gambling and non-gambling sessions of Experiment II produced similar probability preferences, consistency, and intrasitivity, indicating that the *S*s were motivated to make careful decisions under both conditions. The short duration of each session, the individual attention which the *S*s received, and the *S*s' knowledge that they eventually would have to play the bets for money probably combined to make them decide carefully in the non-gambling sessions.

The large increase of simple orderings in Session 6, accompanied by increased consistency and transitivity, indicates that the lengthy, group administered, make believe gambling session bored the *S*s and so induced them to flee the task by adopting quick and easy strategies. Coombs and Pruitt found more simple orderings than we did in Session 6, and our Experiment I produced still more simple orderings than did the Coombs and Pruitt experiment. Apparently in our Experiment II the previous experience with real gambling left a residue of interest in the task which even the boring conditions of Session 6 could not immediately extinguish.

The researcher on decision processes is faced with a choice between obtaining distorted preferences from bored *S*s and making a careful effort to prevent such distortions by motivating his *S*s effectively. Shorter sessions, individual administration, and real gambling all serve to increase motivation.

The bidding method, because its dependent variable is measured on a ratio-scale, is sensitive to extreme likes and dislikes, while methods which obtain only order relationships from *S*s are not. This virtue of the bidding method is illustrated by the discrepancy between the vote count and the mean bid values for the 1/6 and 5/6 negative *EV*-bets, a discrepancy which exists in spite of the fact that both procedures summarize the same data.

Since the bidding method provides powerful ratio-scale data; since, when this data is reduced to order relationships, it is similar to the data obtained by paired-comparison methods; since the method has the advantage of far greater economy than any ordinal method except complete ranking; and since the *S*s like and are well motivated by bidding procedures, we strongly recommend their future use.

SUMMARY

Review of the findings of previous experiments by Coombs and Pruitt and by Edwards indicates that experiments on processes of decision which use group sessions, imaginary choices, and check-off responses produce more orderly data and simpler laws of preference than experiments in which individual *S*s face real bets for real money. In particular, boring conditions typically produce single-peaked preference-functions, while motivating conditions typically produce multiple-peaked preference-functions. The hypothesis is here examined that the *S*s in imaginary choice-experiments are often so bored that they adopt the simplest available strategy permitting choice without thought. Another purpose of the experiment was to evaluate a single stimulus bidding-method which turned out to be an economical and powerful alternative to the method of paired-comparisons.

Experiment I replicated the general conditions of the Coombs-Pruitt experiment, but used an even more boring task: rank-ordering sets of four bets. The main finding was that 70% of all orderings were single-peaked.

In Experiment II, seven men served as *S*s for six sessions. In the first three sessions, they made paired-comparison choices among bets, playing no bets in Session 1, one-third of the chosen bets in Session 2, and all the chosen bets in Session 3. The single stimulus bidding method used in Sessions 4 and 5 required each *S* to state how much money he would be willing to pay in order to play each bet; the task was presented as a competitive bidding situation. The *S*s worked individually in the first five sessions; in the final session they sat as a group and indicated paired-comparison choices between bets by making check marks in a test-booklet. No bets were played in this session.

The stimuli were 15 two-outcome bets with probabilities ranging from 1/6 to 5/6, displayed as the roll of one die. Expected value-levels were zero, +75¢ and −75¢; variance was held constant.

Data-analyses examined consistency, transitivity, probability preferences, and the occurrence of the simplest kinds of probability-orderings across sessions.

Results were mostly similar across the first five sessions. In Session 6,

increases in consistency and transitivity accompanied shifts and simplifications of probability orderings, suggesting that bored *S*s were attempting to ease their task. This is consistent with findings in the Coombs-Pruitt experiment and in Experiment II, in which preferences were even more orderly under even more tedious experimental conditions. Shorter experimental sessions, individual administration, and real gambling aid in motivating the *S*s and thus help to prevent boredom-induced distortions in preferences.

The bidding method is more economical than paired-comparisons and related methods of detecting preference; it provides powerful ratio-scale data and is sensitive to extreme likes and dislikes. When bidding data are reduced to ordered relationships, the orderings are similar to those obtained by the method of paired-comparison. We strongly recommend future use of the bidding method.

TWO DIFFERENT AFTER-EFFECTS OF EXPOSURE TO VISUAL TILTS

By RICARDO B. MORANT and JUDITH RICH HARRIS, Brandeis University

After inspection of a line tilted slightly off vertical, an objectively vertical line appears tilted in the opposite direction. For it to appear vertical again, it must be slightly tilted in the same direction as the first (inspection) line. Such a change in the apparent position of vertical or horizontal lines after exposure to tilted visual configurations is referred to here as a "tilt after-effect."

Gibson and Gibson and Radner characterized the effect as the result of adaptation toward a norm.[1] They maintained that during inspection the tilted line gradually tends to appear less tilted. A line slightly tilted from vertical tends to appear more vertical; one slightly tilted from horizontal tends to appear more horizontal. As a consequence of this, objectively vertical and horizontal lines which are then viewed appear tilted in the opposite direction. If the inspection-line is midway between the vertical and horizontal axes, at a 45° tilt, no normalization to either axis should occur. Therefore, there should be no after-effect.

One particularly interesting finding reported by Gibson and Radner was some interdependence between the horizontal and vertical axes. For example, after exposure to a tilt off vertical, a small after-effect was shown on a horizontal test-line. Gibson and Radner called this inter-axial after-effect the "indirect effect."[2]

Köhler and Wallach disagreed with Gibson's interpretation of tilt after-effects.[3] They postulated that the prolonged inspection of *any* visual configuration leads to localized "satiation" of the stimulated area of the visual cortex. Other figures then tend to be displaced from this satiated region, causing the perceptual distortions they called "figural after-effects." This theory can account for much of Gibson's data—particularly those characteristic of the effect which Gibson called "areal

* Received for publication November 5, 1963. This investigation was supported in part by Research Grant M-3658 from the National Institute of Mental Health.

[1] J. J. Gibson, Adaptation, after-effect and contrast in the perception of tilted lines: II. Simultaneous contrast and the areal restriction of the after-effect, *J. exp. Psychol.*, 20, 1937, 553-569; J. J. Gibson and Minnie Radner, I. Quantitative studies, *ibid.*, 20, 1937, 453-467.

[2] Gibson and Radner, *op. cit.*, 463-466. The controversy over the existence of this indirect effect has been summarized in M. M. Taylor, Figural after-effects: A psychophysical theory of the displacement effect, *Canad. J. Psychol.*, 16, 1962, 247-277. The experimental conditions under which the effect can be demonstrated, and a description of how it differs from the so-called "Wertheimer effect," may be found in R. B. Morant and Mildred Mistovich, Tilt after-effects between the vertical and horizontal axes. *Percept. mot. Skills*, 10, 1960, 75-81.

[3] Wolfgang Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. phil. Soc.*, 88, 1944, 269-357.

restriction."[4] Köhler and Wallach's explanation could not, however, account for the indirect effect. Under the conditions in which Gibson obtained an indirect effect, their satiation-theory would predict either no effect at all, or an effect in the opposite direction. This prediction is clarified in Fig. 1.

Fig. 1 shows the after-effects on a vertical test-line that should occur if only a satiation-like process were operative. The abscissa represents the tilt at which the inspection-line is set, from vertical (0°) through horizontal (90°). The ordinate shows the direction and relative amount that an *S* who had inspected the tilted line would have to adjust a vertical line

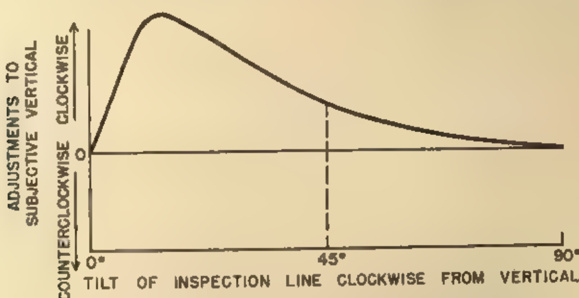

FIG. 1. HYPOTHESIZED EFFECTS OF SATIATION-LIKE PROCESS.

in order for it to appear vertical to him. Of course, the absolute magnitudes or the precise shape of the function are not predictable from the theory. The curve is meant only to illustrate the two predictions which derive from satiation-theory: (1) that the effect should reach a maximum when the inspection-line forms a small acute angle with the test-line, and then gradually decline to zero at greater angles; (2) that the effect should always be in the same direction, regardless of how much the inspection-line is tilted (between 0° and 90°).

Fig. 2 illustrates the type of function a normalization-like process should generate. Again, the expected effect on a vertical test-line is depicted. The predictions are: (1) the effect should reach a maximum when the inspection-line is at some small degree of tilt, and then decline to zero at a 45° tilt; (2) with inspection-tilts greater than 45° an effect should again appear, but in the opposite direction. This is because a line tilted more than 45° off vertical should tend to become more horizontal during inspection, rather than more vertical. If the vertical and horizontal axes of space are dependent on each other—if normalization consists of

[4] Gibson, *op. cit.*, 557-561.

a general reorientation of visual space—then this adaptation toward horizontal would affect a vertical test-line as much as adaptation toward vertical, but in the opposite direction.

In keeping with Gibson's terminology, we shall use the term "direct effect" to refer to the effect on a vertical test-line of normalization toward vertical (or on a horizontal test-line of normalization toward horizontal). The "indirect effect," then, is the effect on a vertical test-line of normalization toward horizontal (or on a horizontal test-line of normalization toward vertical). We shall call the degree of tilt at which the direction of the after-effect changes and no after-effect occurs the "crossover point."

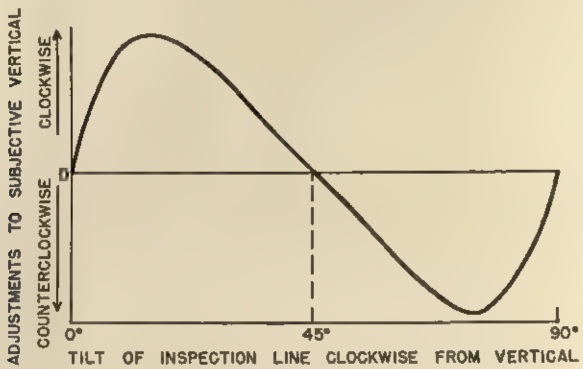

FIG. 2. HYPOTHESIZED EFFECTS OF NORMALIZATION-LIKE PROCESS

It should be stressed once more that both curves are somewhat arbitrary. The absolute magnitudes of the effects as a function of the degree of tilt of the inspection-line cannot be precisely specified from the theories.

Several facts led us to believe that neither process by itself is sufficient to account for tilt after-effects. On the one hand, we have noted that the satiation-like process cannot account for the indirect effect reported by Gibson. Nor can it account for the apparent righting of a tilted line which occurs during its inspection.[5]

On the other hand, two observations raise difficulties for the normalization-like process. The first of these is Köhler and Wallach's demonstration that tilt after-effects occur even when the inspection-line itself is vertical and the test-line is tilted—*i.e.* tilt after-effects occur after inspecting lines already at a norm.[6] The second is Gibson and Radner's and Morant and

[5] Gibson and Radner, *op. cit.,* 454; Morant and Mistovich, *op. cit.,* 80.
[6] Köhler and Wallach, *op. cit.,* 309-311.

Mistovich's observation that the indirect effect is smaller than the direct effect.[7] As we noted above, the reasonable assumption that the axes of space are rigidly linked together implies that the two effects should be equal in magnitude. This difficulty has been discussed by Morant and Mistovich.[8] They suggested that the observed difference in magnitude can be accounted for by postulating that "two factors might be operative with both summating when the test-figure is oriented towards the same axis as the inspection-figure, and in opposition when the test-figure is oriented

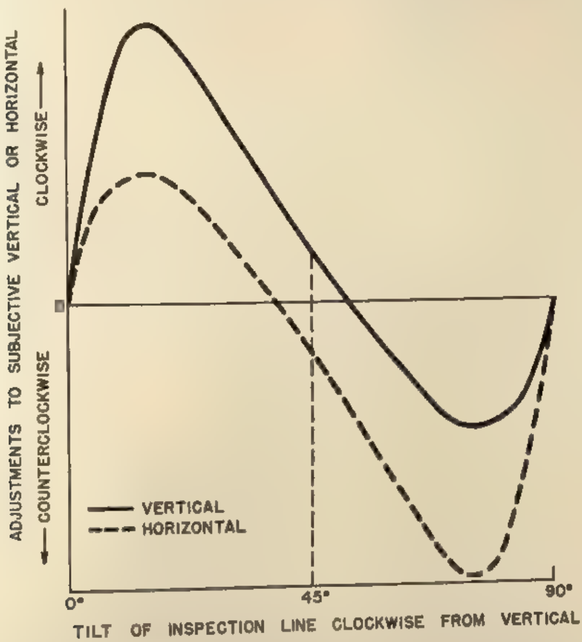

FIG. 3. ALGEBRAIC SUMMATION OF HYPOTHESIZED EFFECTS OF SATIATION-LIKE AND NORMALIZATION-LIKE PROCESSES

toward a different axis." If this were true, the crossover point would *not* be at 45°. Fig. 3 illustrates this prediction.

The solid curve is the algebraic summation of the two curves presented in Figs. 1 and 2. It shows, as a function of the degree of tilt of the inspection-line, the after-effect on a vertical test-line that would occur if both processes were operative. The dotted line shows the curve that

[7] Gibson and Radner, *op. cit.*, 466; Morant and Mistovich, *op. cit.*, 79.
[8] Morant and Mistovich, *op. cit.*, 80.

would result for a horizontal test-line.[9] Note that in both cases the indirect effect is smaller than the direct effect. The crossover point is *before* 45° for the horizontal test-line, and *beyond* 45° for the vertical test-line.

Our first experiment was designed to determine the empirical functions relating after-effects on vertical and horizontal test-lines to tilt of the inspection-line. According to our prediction, these functions should resemble those in Fig. 3.

In a second experiment, we attempted to factor out the effect of the normalization-like process, in hopes that the remaining after-effect would have the characteristics of the satiation-like process. Morant and Mikaelian have already shown that tilt after-effects are not strictly localized, that is, that they are not restricted to the area of the visual field corresponding to the location of the inspection-line.[10] They report, however, that the intra-field after-effects (inspection- and test-lines on the *same* side of the visual field) are larger than inter-field after-effects (inspection-line on one side of the visual field and test-line on the other side). Again, this discrepancy may be due to the fact that two after-effects summate in the intra-field conditions, but not in the inter-field conditions. If we assume that the effects of the normalization-like process are not localized, then, by utilizing the inter-field procedure ("split-field technique"), we should be able to factor them out, leaving only the after-effects of the localized satiation-like process. This prediction will become clearer when the design and results of the second experiment are considered.

METHOD

Experiment I. The tilt of the inspection-line was the independent variable. The dependent variable was *S*s' adjustments of a line to apparent vertical or horizontal. Seven undergraduate students, who were not aware of the purpose of the experiment, served as *S*s. All experimenting was carried out in a totally dark room. A lumiline bulb, taped to allow a line of light ⅛" by 11½" to be visible, served as both the inspection- and test-figures. The bulb was pivoted at its center, and could be rotated in the plane perpendicular to the *S*s' line of sight. A protractor which moved with the bulb, and a stationary pointer, allowed the tilt of the line of light to be read to the nearest quarter degree. Two smoothly-turning knobs permitted either *S* or *E* to adjust the line, through a system of gears and belts. *S* was

[9] The adaptation-like curve for a horizontal test-line is the same as that for a vertical test-line; the satiation-like curve is the right-left and up-down reversal. Accordingly, the resultant curve for the horizontal test-line is the up-down and right-left reversal of the curve for the vertical test-line.

[10] R. B. Morant and H. H. Mikaelian, Inter-field tilt after-effects, *Percept. mot. Skills*, 10, 1960, 95-98.

seated 7.5 ft. from the apparatus, with a fixation-point in the middle of the illuminated line at eye-level. A biteboard was used to reduce head movements.

The testing procedure was as follows. With *S*'s eyes closed, the line was set at a given inspection-tilt. *S* opened his eyes and fixated the point in the center of the line for 1 min. He then closed his eyes, and *E* set the line to objective vertical. Then, 7.5 sec. after the end of the inspection-period, *S* opened his eyes and so adjusted the line that it appeared vertical, and again closed his eyes. The line was reset to the same inspection-inclination, and another 1-min. inspection-period followed. *S* then closed his eyes while *E* set the line to objective horizontal. Again, 7.5 sec. after the end of this second inspection-period, *S* opened his eyes and so adjusted the line that it appeared horizontal. This sequence was repeated at least 4 times by each of the 7 *S*s for each angle of inclination of the inspection-line. A rest-period of 5 min. or more intervened between the end of a test and the beginning of the next inspection-condition. Of the 7 *S*s, 4 were tested with 11 angles of inclination of the inspection line (20° to 70° clockwise from vertical in 5° steps), and 3 were tested with 23 angles of inclination (10° counterclockwise from vertical through 10° clockwise from horizontal in 5° steps). Four different mixed orders of presentation of the various inspection-conditions were used.

Just previous to each inspection-condition, four control readings (two each at vertical and horizontal) were taken for each *S*. These consisted of adjusting the test-line to apparent vertical or horizontal without the previous 1 min. exposure to the inspection-line. All data reported in this paper are in terms of differences between test-measurements and the appropriate control measurements.

Experiment II. The tilt of the inspection-line again served as the independent variable. *S*'s task was so to adjust a line that it appeared *parallel* to an objectively vertical line. Five undergraduate students served as *S*s.

The apparatus and test-procedures for Experiment II were like those of Experiment I, except for the following modifications. Two lumiline bulbs were symmetrically placed on either side of an illuminated fixation-point. The midpoints of the lines were 22″ apart and at the same height. During inspection only the line to the left of the fixation-point was visible. *S* was instructed to keep his eyes on the fixation-point during all inspection- and test-periods. After a 1-min. exposure to the tilted line, *S* closed his eyes while *E* oriented the left line to objective vertical, and turned on the right line. (The line on the right was always objectively vertical.) *S* then opened his eyes and so adjusted the left line that it appeared parallel to the one on the right. This procedure was repeated 5 times in succession for each of 9 angles of tilt of the inspection-line (15° counterclockwise from vertical through 15° clockwise from horizontal in 15° steps) presented in a mixed order. On a later day, the whole series was repeated in reverse order.

Just previous to each inspection-condition, each *S* made five control adjustments. These differed from the test-adjustments only in that no inspection-period preceded them.

RESULTS

Experiment I. The results of the first experiment are summarized in Fig. 4. It will be noted that the direct effect is greater than the indirect for both the vertical and horizontal test-conditions. As predicted, the

crossover point is beyond 45° for the vertical test-conditions and before 45° for the horizontal test-conditions. One-tailed t-tests were used to test the null hypothesis that the mean adjustments did not differ significantly from zero (t-tests were computed for only those conditions where $N = 7$).[11]

Adjustments to horizontal proved more variable than adjustments to vertical, and consequently fewer means differed significantly from zero.

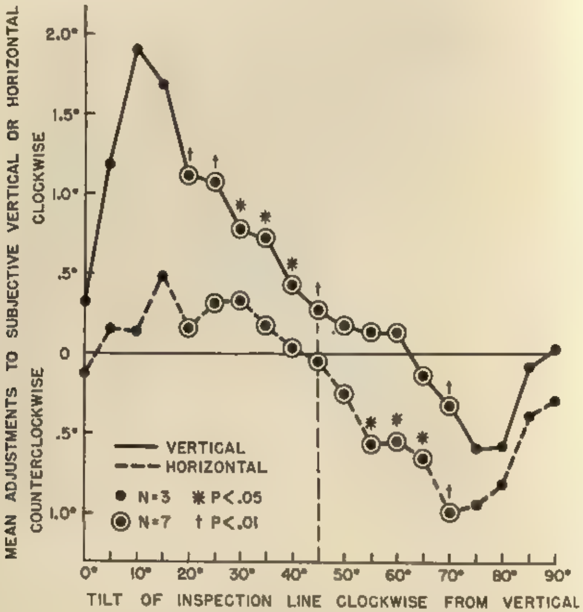

FIG. 4. OBTAINED EFFECTS RELATING ADJUSTMENTS TO INSPECTION TILTS: EXPERIMENT I

Some *S*s mentioned that they had more difficulty making the horizontal than the vertical settings.

Experiment II. The results of the second experiment are summarized in Fig. 5. Again, one-tailed t-tests were used to determine whether the mean adjustments differed significantly from zero. Note that there is no crossover point—the after-effects are all in the same direction of tilt as

[11] To facilitate the comparison of the empirical with the hypothesized curves, only the data for the 0° to 90° inspection-conditions are plotted. The mean results for the other conditions follow: (1) Inspection-line (I) at 10° ccw: Test on Verticle. (V) $= -1.5°$, Test on Horizontal (H) $= 0.0°$. (2) I at 5° ccw; Test $V = -1.0°$, Test $H = -0.4$ (3) I at 95° cw; Test $V = +0.1°$; Test $H = +1.3°$. (4) I at 100° cw: Test $V = +0.2°$, Test $H = +1.0°$. These results are essentially in accord with those plotted.

the inspection-line. As predicted, these results are in accord with the hypothesized effects of the satiation-like process.[12]

We assume that the inspection of the tilted line in the left half of the visual field gives rise to both normalization-like and satiation-like processes. Both of these processes affect the apparent orientation of the test-line on the left. According to this view, however, only the effects of the

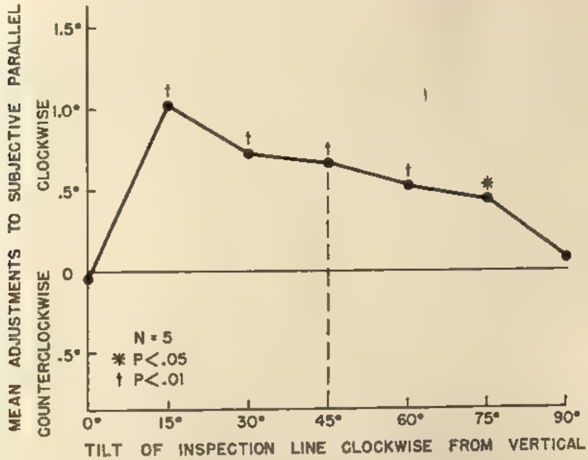

FIG. 5. OBTAINED EFFECTS RELATING ADJUSTMENTS TO INSPECTION-TILTS: EXPERIMENT II

normalization-like process transfer to the other side of the visual field and affect the apparent orientation of the right test-line as well. Since *S*'s task is so to adjust the left test-line that it appears parallel to the line on the right, any process that equally affects both lines will be factored out. Thus, *S*'s adjustments should show only the effects of the satiation-like process, which we assume is confined to the area stimulated.[13]

---

[12] Again to facilitate the comparison of the empirical with the hypothesized curve, only the data from the 0° to 90° inspection-conditions are plotted. The results for the other two conditions follow: (1) *I* at 15° ccw = − 1.0°. (2) *I* at 105° cw = + 0.1°. Once more, these results are essentially in accord with those plotted.

[13] The assumption that the normalization-like effect is not restricted to the locus of the visual field corresponding to the part of the retina stimulated is critical to this argument. The demonstrations purporting to show such localization (Gibson and Radner, *op. cit.*, 556-561; Richard Held, Localized normalization of tilted lines, this JOURNAL, 76, 1963, 147-148) can be shown to be either not directly relevant to the question of localization (Morant and Mikaelian, *op. cit.*, 97-98), or due to artifacts of the experimental procedure used (Morant, Is the normalization of tilted lines localized?, to be published).

Discussion

We have already noted the reasons which led us to assume that two processes must be postulated to account for tilt after-effects. The results of the two experiments reported here further attest to the usefulness of this assumption. It should be stressed, however, that our identification of the two processes as "satiation-like" and "normalization-like" should not be taken as a commitment to the specific theories of Köhler and Wallach or Gibson. For example, Osgood and Heyer have proposed a theory designed to account for the same observations covered by the Köhler-Wallach theory.[14] Werner and Wapner have interpreted some of Gibson's observations within the framework of their sensory-tonic field-theory.[15] More recently, Taylor has proposed a psychophysical theory to account for various examples of both the Köhler-Wallach and Gibson effects.[16]

We are not particularly concerned with any specific one or two of the theories—the grounds for preferring one to another are not yet clear—but only with the effects they predict. Thus by "satiation-like process" we have meant a tendency for a new contour to be displaced from the retinal location of an old one. By "normalization-like process" we have meant a tendency for the position in visual space which is considered to be vertical or horizontal to shift after exposure to visual tilts.

Summary

Our results indicate that two factors produce tilt after-effects: a localized process resembling Köhler and Wallach's "satiation," and a nonlocalized process resembling Gibson's "normalization." When the inspection- and test-lines are in the same part of the visual field, these factors summate algebraically. They add to each other when the inspection-line is tilted less than 45° from a vertical or horizontal test-line, and they act in opposition at greater tilts. If a comparison figure in a different part of the visual field is used, only the satiation-like process can be shown.

---

[14] C. E. Osgood and A. W. Heyer, A new interpretation of figural after-effects, *Psychol. Rev.*, 59, 1952, 98-118.

[15] Heinz Werner and Seymour Wapner, Experiments on sensory-tonic field theory of perception: IV. Effect of initial position of a rod on apparent verticality, *J. exp. Psychol.*, 43, 1952, 68-74.

[16] Taylor, *op. cit.*, 247-277. For a review of other theories proposed to explain figural and normalization after-effects, see Peter McEwen, *Figural After-Effects*, 1948, 42-97; Morigi Sagara and Tadasu Oyama, Experimental studies on figural after-effects in Japan, *Psychol. Bull.*, 54, 1957, 327-338.

MILD STRESS AND PROBLEM-SOLVING

By WILBERT S. RAY, Bethany College

This paper reports three experiments to test the assumption that stress interferes with problem-solving. Two of the experiments supported the hypothesis, while the third was an unsuccessful replication of a previous experiment. The stress used in each case was rather mild. It was an experimentally-induced drive rather than the generalized drive inferred from manifest-anxiety scales.[1]

Problems exist only in situations in which the correct response is not the first to appear, is not the dominant response. It has been pointed out, *e.g.* by Maltzman and Taylor, that Hull's multiplicative drive-theory implies that an increase in drive-strength will increase the reaction-potential of the already dominant response more than the reaction-potential of the correct but weaker response.[2] It will therefore be more difficult for the organism to produce the correct response as drive-strength increases. The effect will be greater with more complex problems with their greater difference between the dominant and non-dominant responses. Hull has applied this hypothesis, that increasing drive will increase the differential, to one specific sort of problem-solving—the *Umweg* problem.[3]

REPETITION OF AN EXPERIMENT ON PERSONAL AND IMPERSONAL CONDITIONS

Marks reports an experiment in which, among other variables, a "personal" situation lowered the number of successes in problem-solving as compared with an "impersonal" situation.[4] The *S*s learned to find square roots by using a calculator and the Marchant Calculating Machine Company's table of "Square Root Divisors." In this experiment, which was repeated here, two groups of *S*s used a typed copy of thirteen factors from

* Received for publication June 14, 1962. The experiments reported here were partially supported by the Office of Naval Research under Contract Nonr 2315(oo).

[1] The experiments add to existing information about problem-solving. See C. P. Duncan, Recent research on human problem solving, *Psychol. Bull.*, 56, 1959, 397-429; W. S. Ray, Complex tasks for use in human problem solving research, *Psychol. Bull.*, 52, 1955, 134-149; and B. J. Underwood, An orientation for research on thinking, *Psychol. Rev.*, 59, 1952, 209-220.

[2] Irving Maltzman, Thinking: from a behavioristic point of view, *Psychol. Rev.*, 62, 1955, 275-286; J. A. Taylor, Drive theory and manifest anxiety, *Psychol. Bull.*, 53, 1956, 303-320; C. L. Hull, *Principles of Behavior*, 1943, 238-240; *A Behavior System*, 1952, 265-268.

[3] Hull, *op. cit.*, 1952, Theorem 101.

[4] M. R. Marks, Problem solving as a function of the situation, *J. exp. Psychol.*, 41, 1951, 74-80.

the Marchant table, but with two entries falsified. The two were different from the other eleven numbers in the second place of the seven-digit factors.

After four practice exercises, the *Ss*, in the Personal group, work four more problems, including two which use the falsified factors. The *S*'s attention is then called to the two erroneous roots, and he is asked to ascertain the source of the errors. An *S*, in the impersonal group solves the four practice exercises but does not work the four test-problems. Instead, he is told that John Jones has solved four more problems, but produced two erroneous answers. He was asked what he would do if he were in Jones' place and had to correct the errors. The criterion of his success was whether he discovered the source of Jones' errors.

Marks predicts and reports a much higher proportion of success among the *Ss* in the impersonal situation than in the personal condition. (See our Table I.) On the other hand, in the personal situation, *S* actually used the incorrect factors, and this should later help him to locate the source of the errors when errors appear. The prediction on this basis would be that in the Personal condition the *S* who

TABLE I

NUMBERS OF SUCCESSES AND FAILURES IN THE CALCULATOR PROBLEM

	Personal group	Impersonal group	Added group	Chi-square	*P*
1951 experiment					
N succeeding	5	16			
N failing	25	14		8.69	<.01
Present experiment					
N succeeding	4	2	9		
N failing	8	10	3	8.92	<.01

had worked the four test-problems would have a higher proportion of success than did those in the Impersonal condition who did not work them. This prediction is opposed to that based on the motivational factor.

The present experiment repeated these two conditions of Marks. We also, however, added a third group, which worked the four practice-problems *and* the four test-problems, and were then told that Jones had had the same trouble, and asked "What would you do if you were Jones?" Here the number of successes should be raised by the impersonal condition, and by the practice with the incorrect entries. Assuming the correctness of Marks' reported facilitation by the impersonal situation, this third treatment should raise the number of successes still higher.

Subjects. The *Ss* used here were 36 students from a course in General Psychology who participated as part of their assigned course work. Two additional *Ss* were discarded for failure to learn the square-root procedure

Results. The results of the present experiment are presented in Table I. For comparison, those of the previous experiment are also shown.[5]

[5] The Marks figures are from his Group I and Group II, Table I, recalculated. See Marks, *op. cit.*, 4, 80.

The present results show a significant difference among experimental treatments, but the difference is between the added group and the two replicated groups. Our Personal and Impersonal groups do not differ significantly from each other. This experiment does not support the finding of the previous experiment that personal involvement lowered the probability of success in problem solving.

Summary. The lack of confirmation of the previous experiments suggests that somehow we did not in fact use the same procedures that Marks has used. We are unable, however, to find the difference.

We predicted above that actual use of the falsified factors in the test-problems would raise the number of successes in the personal group (as well as in our added group), but this did not happen. Perhaps the inhibition due to personal involvement was counterbalanced by the facilitation due to the practice with the incorrect factors.

Set for Speed as an Inhibitor of Problem-Solving

This is a repetition of an experiment on inhibition accompanying a set for speed performed by the writer's classes, in Experimental Psychology in 1957.[6]

There is an experimental basis for set for speed as an independent variable in the work of Weiss, who found an inhibitory effect from a set for speed in a situation that is at least much like problem-solving.[7] The problem employed was a simplified form of the puzzle, "The Tower of Hanoi"[8]—also known as the "Tower of Brahma"—in which *S* transfers a tower of disks from one peg to a vacant peg, moving one disk at a time without placing a larger disk on top of a smaller, and using a third peg as a way-station. The shortest solution of the problem in number of moves being 2^n-1 disks, an n of only four disks was used *S* made the transfer only once each.

Conditions. In the experimental condition the *Ss* were requested to work as fast as possible, the request being emphasized by ostentatiously starting a stopwatch with the signal to begin. The control group did not receive these instructions, or see the stopwatch.

Subjects. The *Ss* were 32 students from a class in General Psychology who participated in the experiment as part of their assigned course work.

Results. All *Ss* solved the problem. Figures for the dependent variables are shown in Table II. The speed-group made more moves, at a statistically

[6] P. A. May, I. A. Mallas, E. E. Ekas, P. S. Chase, Jr., R. S. Baca, L. M. Riley, and W. S. Ray, Set for speed as a variable in problem-solving. *Proc. West Va. Acad. Sci.*, 29-30, 1957-58, 98-99. E. F. Silverg administered the experimental task to the *Ss* both in the present experiment and in the repetition of the Marks experiment; and made valuable suggestions about the work.

[7] R. L. Weiss, The influence of 'set for speed' on 'learning without awareness,' this Journal, 68, 1955, 425-431.

[8] For a description and history of this puzzle see Martin Gardner, *Mathematical Puzzles and Diversions,* 1959, 57 ff.

significant level, but showed no difference in the total amount of time needed for the problem. They worked faster in the sense of using less time per move.

Heterogeneity of variance was unexpectedly found in the previous experiment (May), and since it was found there it was predicted here. The size of the variance in the control groups of the two experiments was about the same, but in the experimental groups the amount decreased markedly, presumably because of more careful control of the experimental conditions. This decreased the difference between the two groups in this present experiment, but the difference is still significant. Since differences in number of moves and of variance were predicted, one-tailed tests are used.

TABLE II

THE EFFECTS OF A SET FOR SPEED
(16 *S*s in each group)

	Speed group	Non-speed group		P*
Mean no. of moves	37.75	24.56	$t = 2.37$	.02
Mean no. of sec.	172.06	173.31		
Variance of moves	347.19	116.00	$F = 2.99$	.03

* Tests are one-tailed

The minimal number of moves necessary to re-stack four disks is 15. If we subtract 15 from each of the means for moves in Table II, we have 22.75 errors in the experimental and 9.56 errors in the control condition.

Discussion. An increase in variance in an experimental group has also been reported by Anderson, who found it when *S*s were required to shift back and forth between two problems at 3-min. intervals.[9] She describes it by saying that "the shift-procedure enhanced differences between individuals." A similar statement applies to our results.

Weiss' report of impairment due to a set for speed has been mentioned above. Bruner, Goodnow, and Austin give somewhat similar results in a concept-attainment task, "time pressure has a relatively small deleterious effect on the success of focusing, but a major (deleterious) effect on the success of scanning."[10]

Our major hypothesis is that stress will show an inhibitory effect on problem-solving, but the results reported and quoted here cannot be interpreted to mean that asking an individual to hurry will necessarily interfere with his work. The heterogeneity of variance suggests that the effect may occur only in some persons. The Bruner, Goodnow, and Austin study

[9] S. B. Anderson, Shift in problem solving, *NRL Memo Report 458.* Washington: Naval Research Laboratory, 1955.

[10] J. S. Bruner, J. J. Goodnow, and G. A. Austin, *A Study of Thinking,* 1956, 147.

suggests that the inhibition appears in some conditions and not in others. We may, then, assume that a set for speed perhaps produces stress in some persons under some conditions, and suggest that the impairment of the problem-solving appears only when the stress appears.

Inhibition Produced by Failure on a Preceding Problem

A third sort of stress which may be easily produced in mild degrees in the laboratory is frustration. It was assumed that failure to solve a problem would produce frustration which would inhibit a following problem, especially if the problems were very similar to each other. The failure was produced very simply by giving *S* an insoluble problem.

Conditions. The experimental task was a row of 46 digits mimeographed across a sheet of paper, containing 10 examples of sequences of three consecutive digits (*e.g.*, 3, 4, 5 or 5, 6, 7), with an *X* below the third digit in each of these sequences. This task was used in a previous study.[11] *S* is requested to discover "What

TABLE III

Number Successful in Solving Numerical Problem after Frustration and No-Frustration Conditions

	Problem first (No frustration)	Problem second (Frustration)
N successful	34	23
N unsuccessful	35	48

Chi-squared = 4.13; $P = .02$

is it in the digits that tells where to put an *X*?" and then to put *X*s in the proper places in a "test" set of 14 digits at the bottom of the page. Time allowed was 12 min.

The frustrating (failure) task was a similar set of digits and *X*s, but with no patterns among them. *S* spent 12 min. on this task also.

The two problems were assembled under a cover sheet, with the insoluble task as the second sheet and the real task as the third sheet for half the booklets, and with the two task-sheets reversed for the other half. These booklets were alternated in a pile and passed out to the *Ss*, who were assembled in one room. It was assumed that this would insure a random assignment of treatments to *S*.

Subjects. The *Ss* were the members of the three sections of a course in Elementary Psychology given in an evening college.[12]

Results. The results of the treatments are shown in Table III. Forty-nine per cent of the *Ss* in the control group were successful, but when

---

[11] W. S. Ray, Generalization among meaningful relations in problem solving, this Journal, 71, 1958, 737-741.

[12] My thanks are due Professors Andrew Leitch and John Korzi of Ohio University, Martins Ferry Branch, for supplying *Ss*.

the insoluble problem was presented first, the number successful fell to 32%. Chi-square was 4.13. Since it was predicted on the basis of the hypothesis about the effects of increased drive, and on the basis of the set-for-speed experiment, that the frustration would inhibit problem-solving, a one-tailed test is applicable, and thus we get a probability of approximately 0.02.

Discussion. This same experiment was tried with a class in elementary psychology at Bethany College some seven months previous to the trial described here. With 110 students divided equally between the treatments, 51% were successful in the control condition, and 38% succeeded in the frustration condition. This gave a non-significant Chi-square of 1.80. The difference between the two experiments was that in the Bethany situation the frustration- and work-periods were 7, rather than 12 min. in duration. After the first trial of the hypothesis we assumed that a longer failure period would produce more frustration and the change from 7 to 12 min. of failure did produce more failures in the succeeding experiments.

Roberts reports results which agree with those reported here.[13] His task was a modification of the "Twenty Questions" task which has previously been used in problem-solving research.[14] Roberts counted the number of clues each *S* received before identifying an object. In one condition the *S*s had more failures in a preceding task than in the other condition, and the mean numbers of clues necessary were, respectively, 18.88 and 16.93. (These figures recalculated from his Table 5.) The analysis of variance showed an *F*-ratio of 3.50, which provides a one-tailed significance level of 0.05.

General Discussion

The hypothesis offered above says that increase in drive strength will produce increasing inhibition of problem-solving, and that the inhibition will be greater with more complex problems. The inhibition will not start with zero drive strength, since there must be some minimal amount of motivation present to produce any work on the problem. That latter statement is reminiscent of analogous statements about the relation of drive strength to degree of learning. Deese concludes that associative factors in learning are not influenced by level of motivation, but "If an animal is to learn, it must be motivated at least enough to be goaded into activity."[15]

The most popular current hypothesis concerning the relation of motivation level to problem-solving proficiency is that given, for example, by Krech and Crutchfield.[16] They say, "As the degree of motivation increases from zero, the problem solving efficiency first increases and then decreases." In support of this function, they quote Birch,

---

[13] J. S. Roberts, Jr. Information seeking in sequential decision making as dependent upon test anxiety and upon prior success or failure in problem solving, in D .W. Taylor (ed.), *Experiments on Decision Making and Other Studies, Tech. Rep. No. 6,* (NR 150-166), New Haven, 1960, 26-48.

[14] D. W. Taylor and W. L. Faust, Twenty questions: Efficiency in problem solving as a function of size of group. *J. exp. Psychol.,* 44, 1952, 360-368.

[15] James Deese, *The Psychology of Learning,* 1958, 107.

[16] David Krech and R. S. Crutchfield, *Elements of Psychology,* 1958, 392.

When motivation is very low the animals are easily diverted from the problem. . . . Under conditions of very intense motivation, the animals concentrated upon the goal to the relative exclusion of other features of the situation which were essential to the solution of the problem. . . . Those animals who worked . . . under intermediate conditions of motivational intensity . . . were not . . . incapable of responding to other relevant features of the problem situation.[17]

Birch's statement can be recast into the form of ours without losing its applicability to his observations. Animals who were not motivated toward the problem did not work on it. Of the animals who were motivated, and who did work on it, those who were the most strongly motivated were the ones who showed the interference.

It is possible to say that the Krech and Crutchfield hypothesis is somewhat similar to ours if their inverted-bow-shaped curve is so distorted that the rising-facilitation half becomes our stipulation that there must be enough motivation present to produce work at the problem.

It is impossible to tell exactly what is meant by 'high' and 'low' drive, and it may be suggested that in the discussion of the interference of high drive with performance and learning 'high' merely means the highest level used in some particular experiment, which gives no information as to the position of the drive level on any general scale of intensity.[18] French's pseudo-fire, and Patrick's shock, cold water, and loud noise seem to be of a nature to produce high degrees of motivation, but Köhler's change of valence with increasing proximity of food is probably not classifiable as "very intense."[19] The set-for-speed and frustration reported herein may be assumed to produce fairly mild intensities. Available evidence, then, supports the position that low degrees of motivation also inhibit problem-solving, and permits the hypothesis that any increase in motivation will interfere with problem-solving processes.

Comparison of the specific drives and the specific experimental tasks reported in the literature suggests that the inhibiting motivation is not necessarily directly related to the task. The hypothesis under discussion refers to the intensive dimension of the motivation rather than to its directional aspect.

Kendler and his associates assume that drive interferes with problem-solving responses that are low in the hierarchy, and not with those higher.[20] This could be restated to say that drive interferes more with responses low in the hierarchies than with those higher, and their results can be interpreted to support this position. Such a statement would correspond to the section of ours that speaks of greater

---

[17] H. G. Birch, The role of motivational factors in insightful problem-solving, *J. comp. Psychol.*, 38, 1945, 295-317, esp. 316.

[18] See, for recent examples, R. E. Clark, The role of drive (time stress) in complex learning: An emphasis on prelearning phenomena, *J. exp. Psychol.*, 63, 1962, 57-61; W. E. Broen, Jr., and L. H. Storms, A reaction potential ceiling and response decrements in complex situations, *Psychol. Rev.*, 68, 1961, 405-415.

[19] J. R. P. French, Jr., Studies in topological and vector psychology: III. Organized and unorganized groups under fear and frustration, *Univ. Iowa Stud. Child Welfare*, 20, 1944, 231-308; J. R. Patrick, Studies in rational behavior and emotional excitement: II. The effect of emotional excitement on rational behavior in human subjects. *J. comp. Psychol.*, 18, 1934, 153-195; Wolfgang Köhler, *Mentality of Apes*, 1931, 14.

[20] H. H. Kendler and T. S. Kendler, Vertical and horizontal processes in problem solving, *Psychol. Rev.*, 69, 1962, 1-16.

interference with more complex problems. Spence, Farber, and McFann's report on competitional and non-competitional paired-associates learning also supports this position on complex tasks.[21]

If the hypothesis that increase in drive is accompanied by increasing inhibition of problem-solving, and that the increase is greater as problem complexity grows, is accepted, then it becomes desirable to attempt to explain the mechanism of such interference. Kendler and Kendler say, "It would be expected that a strong drive would retard problem-solving because it would retard the extinction of the dominant incorrect response."[22] This would apply to both parts of our present hypothesis.

A further, and not incompatible, hypothesis is the generalization that an increase in drive will restrict the range of cue-utilization and thus interfere with at least some sorts of learning and perception, which may easily be extended to problem-solving on the assumption that ignoring of cues will prevent solutions of at least some problems.[23] This would describe the work of the chimpanzees in the quotation from Birch above, and the behavior of Köhler's bitch in an *Umweg* problem of which he says, "When . . . the food . . . was dropped just outside the fence . . . she stood seemingly helpless, as if the very nearness of the object and her concentration upon it . . . blocked the 'idea' of the wide circle around the fence."[24]

Summary

Three experiments are described. One was a repetition of a previously reported experiment in which a personal condition interfered with problem solving as compared with an impersonal condition. The replication failed to find a difference between the conditions. In another experiment a set for speed inhibited problem solving, as did frustration in the third.

It was suggested that, above a minimal level necessary to produce work at the problem, further increase of drive-level would produce increasing inhibition of problem-solving, the effect being greater with complex problems. Our experiments contradict the first half of the inverted-bow-shaped relation between motivation and problem-solving which has been postulated by previous theorists.

---

[21] K. W. Spence, I. E. Farber, and H. H. McFann, The relation of anxiety (drive) level to performance in competitional and noncompetitional paired-associates learning, *J. exp. Psychol.*, 52, 1956, 296-305.

[22] Kendler and Kendler, *op. cit.*, 15.

[23] J. A. Easterbrook, The effect of emotion on cue utilization and the organization of behavior, *Psychol. Rev.*, 66, 1959, 183-201.

[24] Köhler, *op. cit.*, 14. This is the case to which Hull applied his hypothesis of the multiplicative drive. See Hull, *op. cit.*, 1952, 266.

PREVIOUS AND CONCURRENT VISUAL EXPERIENCE AS DETERMINANTS OF PHENOMENAL SHAPE

By DOROTHY DINNERSTEIN, Rutgers University

The results to be described here illustrate an alternative to the generally accepted dichotomy between perception and memory, or between the intrinsic structure of the field and the effects of previous experience. When antecedent and concurrent determinants of visual shape are compared, it becomes clear that certain effects of past experience on present experience correspond so closely with effects of concurrent influences as to make unnecessary the use of separate terms in thinking about them. The matter can instead be formulated as follows: A visual area, like any other stimulus-item, is part of more than one functionally relevant larger unit. Phenomenal shape is governed by, and can be altered by change in any one of, a number of larger structures, some spatially and some temporally extended. When—as often happens—the influence of one such larger structure competes with the influence of another, the outcome depends on their relative strength. This is as true when one of the competing structures is temporal and the other spatial as when both are spatial.

EXPERIMENT I: THE INTERACTION OF SPATIAL AND TEMPORAL STRUCTURES IN DETERMINING THE SHAPE OF A FLAT AREA

Shown in *A* of Fig. 1 is Area X, the visual entity whose context is manipulated in this experiment. In *B* of Fig. 1, the presence of an adjoining area changes Area X from an *el* to a partly hidden member of a pair of squares. An earlier study has shown that while in Fig. *B* Area X is described as square by virtually all *S*s, in Fig. *C* it is *el*-shaped for over half of them.[1] In that study, which dealt with a number of cases of phenomenal overlapping, configurational factors such as similarity, simplicity, and symmetry were stressed. Other writers have pointed to familiarity as an additional factor.[2] Indeed, since such natural forces as gravity and

* Received for publication June 30, 1963.

[1] Dorothy Dinnerstein and Michael Wertheimer, Some determinants of phenomenal overlapping, this JOURNAL, 70, 1957, 21-37. There it was concluded that whether one of two contiguous areas will appear to continue behind the other depends not only on the shapes of both areas, but also on outlying, non-contiguous areas.

[2] Alphonse Chapanis and R. A. McCleary, Interposition as a cue for the perception of relative distance, *J. gen. Psychol.*, 48, 1953, 113-133.

surface-tension guarantee that structurally stable and economical shapes will have been encountered frequently in the past, prior experience even may be regarded as an *alternative* explanatory principle,[3] with learning

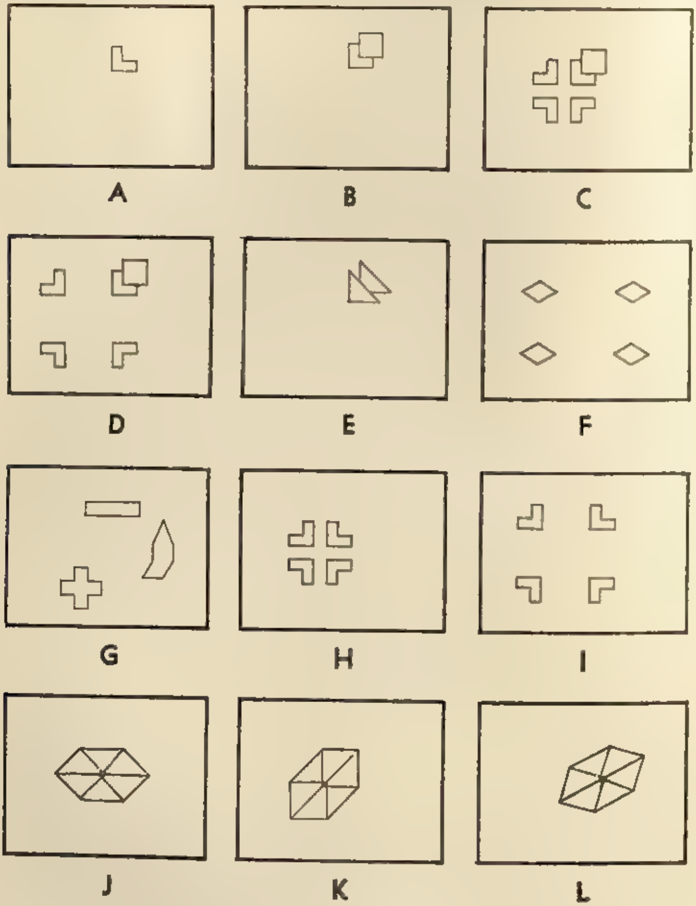

FIG. 1. THE DRAWINGS USED IN THE EXPERIMENTS

perhaps underlying the importance of configurational factors.[4] One can, on the other hand, take the contrary view that configurational factors, operating across time, underlie the interaction between earlier and later

[3] C. B. Zuckerman and Irvin Rock, A reappraisal of the roles of past experience and innate organizing processes in visual perception, *Psychol. Bull.*, 54, 1957, 269-296. Irvin Rock, The present status of Gestalt psychology, in J. G. Peatman, and E. L. Hartley (eds.), *Festschrift for Gardner Murphy*, 1960, 127-128.

[4] J. J. Gibson, *The Perception of the Visual World*, 1950, 142-143.

experience and make learning possible.[5] Thus, we can think of successive experiences as forming temporally extended groups in some ways comparable to such spatially extended groups as the square- and *el*-constellation of Fig. *C*. It then becomes possible to study the roles of prior and concurrent experience by comparing the effects of variations in two of the larger structures of which Area X is part, one spatial and the other temporal.

Experiment IA: Effect of varying a spatial structure. Presumably Area *X* looks more like an *el* in Fig. *C* because the dense nearby group of similar areas can be seen as a symmetrical constellation only if Area *X* is *el*-shaped. If this is true—

TABLE I

EFFECTS OF SPATIAL AND TEMPORAL CONTEXT: EXPERIMENTS IA AND IB

Preceding series	Drawing described	*N*	Percentage of *S*s describing area *X* as		
			Square	*El*	Ambiguous
N	1*B*	23	96	4	0
N	1*C*	32	28	53	19
N	1*D*	56	79	12	9
EL	1*D*	31	42	42	16

that is, if the change from Fig. *B* to Fig. *C* does reflect competition between two structures and not, say, the mere presence of additional visual items—then weakening one of these structures should increase the influence of the other. For this first experiment, the *el*-constellation was weakened, as in Fig. *D*, by moving its members further apart.[6]

As in all the experiments to be described, black line-drawings on 8 × 11-in. white cards were displayed to classroom groups of students in undergraduate psychology. In the present case, Figs. *B, C,* and *D* were shown to three separate groups of Brooklyn College *S*s, preceded in each case by Series *N*, which consisted of Figs. *E, F,* and *G*. Each group saw, in succession, the three members of the (neutral) Series *N*, followed by *one* of the other figures. Each figure was displayed for 10-12 sec., after which the *S*s replied in writing (the reply took 60-90 sec.) to the request "Describe briefly what this is. State what surfaces, if any, appear to go behind other surfaces." The next drawing then was shown.

Results. Descriptions of the Series-*N* figures were examined only to confirm that these drawings were, as expected, seen in a generally uniform way by all the *S*s. The first three rows of Table I show how Area *X* was characterized.[7] In Fig. *B*, *X* was described as one of a pair of squares

[5] Wolfgang Köhler, *Gestalt Psychology*, 1947, Ch. 8, 9, *e.g.* 277-278, 318.

[6] Another experiment, not reported here, gave the same results when the symmetry of the constellation, rather than its density, was reduced.

[7] As the quoted protocols indicate, such categorization generally was a fairly simple matter. Subsequent tests showed inter-judge disagreement in fewer than 5% of cases.

("two overlapping squares," "two squares, one covering the other in part") by nearly all the *S*s. In Fig. *C*, it was so described by less than one-third of the *S*s ($P < 0.001$),[8] while more than half described it as one of a set of *el*s ("four *L*-shaped figures with one square next to one of the *L*-figures," "four *L*-shaped figures in box-formation with the upper-right *L* partly covered by a square").[9] In Fig. *D*, however, Area *X* again was described as a square ("three *L*s," and "one square overlapping another square") by most *S*s; it appeared as an *el* less than one quarter as often as in Fig. *C* ($P < 0.001$).

Experiment IB: Effect of varying a temporal structure. Could the ineffectual constellation-quality of the *el*s of Fig. *D* now be supported by embedding them in turn in some strong larger group of constellations, just as the ineffectual *el*-quality of Area *X* in Fig. *B* was supported by embedding it in the strong constellation of *el*s in Fig. *C*? If structures across time do have functional resemblances to those in space, it should be possible to accomplish this with a temporally extended group of constellations.

In Experiment IA, one of the local *spatial* structures embracing Area *X* was altered, while the immediate temporal context was kept constant. In the present experiment, a local *temporal* structure—the pre-series—was altered, while the spatial context was kept constant. A new group was shown Fig. *D* preceded by Series *El*, which consisted of Figs. *E*, *H*, and *I*.[10]

Results. As the last two rows of Table I show, membership in a series of *el*-constellations (*El*) instead of in a neutral series (*N*) dramatically strengthens the effect of Fig. *D*'s weak *el*-constellation. It triples the percentage of *S*s describing Area *X* as an *el*, and halves the percentage describing it as a square ($P = 0.002$). Table I, then, makes the broad point that certain effects of earlier on later experiences correspond functionally with effects of concurrent experiences on each other. A temporal structure (Series *El*) like a spatial one (Fig. *C*'s *el*-constellation) can, if it is strong enough in relevant respects, fortify one of its members against the requirements of a competing structure.

---

[8] The *P*-values quoted in this paper are based on Fisher's exact test of the raw frequencies of two main types of description. They were calculated with the coöperation of Professor F. Lehman, of the Newark College of Engineering, on an IBM-computer already programmed, in connection with another project, to perform this operation.

[9] These results replicate those of Dinnerstein and Wertheimer (*op. cit.*, 26-27). As before, seeing Area *X* as *el*-shaped apparently did not preclude seeing it continue behind its neighbor. This fact is interesting in connection with the problem of phenomenal overlapping *per se*.

[10] Note that the first member of Series *El* was the same as in Series *N*. It was intended to fulfill the expectation, created by the instructions, that some overlapping areas would appear, leaving *S* freer afterward to describe Area *X* either as *el*-shaped or as square.

Experiment II.: The Interaction of Spatial and Temporal Structures in Determining Depth in an Ambiguous Figure

It should be noted that of the two structures compared in the first experiments, one (the drawing) was a sub-unit of the other (the series). Thus, the two differed not only as to spatial vs. temporal extension but also as to content and internal organization. In the experiment now to be described, the spatial and temporal structures compared were identical in content: as a result, a special characteristic of temporal effects *per se* could be considered.

The same pattern in two different positions is shown in Fig. *J* and *K*. Earlier work has established that virtually all *S*s describe Fig. *K* as a cube, most describe Fig. *J* as a flat hexagonal line-pattern.[11] Is this effect of the orientation of the drawing due to (a) forces of visual organization *per se* which determine whether a two- or a three-dimensional organization of the figure will place it in a more stable relation to the larger spatial structure provided by the edges of the card and the main horizontal-vertical dimensions of the visual field? or is it due to (b) the greater empirical frequency with which rectangular solids in *S*'s past have appeared as resting flat rather than standing on their edges? Stated thus, the question, like the question of the role of familiarity in overlapping, is insoluble. The present experiment was conceived in the view that (a) and (b) do not differ enough to be regarded as true alternatives: the experiential structures in which Fig. *J* and *K* appear can be so varied as to make it clear not only that past experience does indeed influence the phenomenal depth of each of these figures, but also that this influence may closely resemble that of corresponding concurrent experience. The design required that the unequivocally cubic Fig. *K* be replaced with a slightly weaker version of itself which would be described as flat by a minority of *S*s roughly as large as the minority describing Fig. *J* as cubic. In an exploratory study, Fig. *L* was selected to meet this requirement.[12]

Experiment IIA: Effect of varying spatial context. Three separate groups of students at Rutgers University were presented, respectively, either with Fig. *J* alone,

---

[11] Dinnerstein and Wertheimer, *op. cit.*, 30.

[12] The writer is indebted to Anthony Fazio, James DeChesare, and Robert Schoen, who made drawings, served as *E*s, and scored protocols, both for the exploratory study and for an early version of the Experiment II*A*. These preliminary studies revealed effects on phenomenal depth not only of the drawing's position, but also of its size relative to line-thickness and perhaps also to the size of the background-card.

with Fig. *L* alone, or with the two figures side-by-side.[13] In the third condition, position-effects were controlled by placing each figure to the left for half of the *Ss* and to the right for the other half and, since no significant differences appeared, the results for the two subgroups were combined. Exposure- and writing-times were 10 and 90 sec., respectively, when a figure appeared alone; 15 and 150 sec. when the two appeared together. The instructions were to "Describe briefly what this is." Those *Ss* who saw both figures together were asked to describe them on two opposite pages of a booklet in order that the descriptions could be scored independently. Protocols were classified in terms of the presence or absence of depth.

Results. The middle row of Table II shows that Fig. *J*, viewed alone, was described as flat ("a hexagon with lines coming together at center

TABLE II

EFFECTS OF SPATIAL AND TEMPORAL CONTEXT: EXPERIMENTS IIA AND IIB

Condition	Percentage of *Ss* Describing Fig. *J* as				Percentage of *Ss* Describing Fig. *L* as			
	N	flat	cubical	both	*N*	flat	cubical	both*
Simultaneous	44	64	25	11	44	48	43	9
Alone	43	79	21	0	47	26	72	2
Successive	47	38	53	9	43	58	40	2

* Since protocols describing alternative organizations of the figure were more common than in Experiment I, their frequency is noted.

to form a diamond," "a group of triangles that combine to form a six-sided figure," "stop-sign divided into sections") rather than cubical ("rectangle showed in perspective on a white background," "rectangular box—three dimensional picture") by more than a 3:1 ratio, while Fig. *L* yielded the opposite results ($P < 0.001$). The importance for depth of the spatial relation of a figure to the main axes of the visual surround thus is confirmed. The influence of these main axes may, however, be counteracted by that of a conflicting spatial structure. As the top row of Table II shows, Fig. *L*, predominantly cubical alone, is described as flat by half of the *Ss* when accompanied by Fig. *J* ($P = 0.010$). Thus, membership in the pair, whose influence is opposed to that of the spatial framework, proves an effective competing determinant. The predominantly flat Fig. *J*, it should be noted, yields only insignificantly to the presence of Fig. *L*, an asymmetry of influence that will be referred to again below.

The main result to be noted here is that descriptions of Fig. *L* can be

[13] Each experimental group consisted of two small classroom-groups whose data are combined to give the totals of Table II. Differences between the separate groups in no case approached significance.

affected drastically by the presence (concurrent) of Fig. *J*. Experiment II*B* will show that a corresponding, though not identical, effect occurs when the pair-structure competing with the influence of the main axes of the visual field is temporally rather than spatially extended.

Experiment IIB: Effect of varying temporal context. The two groups of *S*s who had, for the purposes of Experiment IIA, seen and described Fig. *J* or Fig. *L* alone immediately afterward were shown and asked to describe the other figure; the interval between the two figures was the 90-sec. period allowed for describing the first, The *S*s turned to a new page in their answer-booklets before seeing and describing the second figure.[14]

Results. The bottom two rows of Table II show for each of these figures the influence of previous experience of the other. Fig. *L*, prevailingly cubical in isolation, was described as flat by more than half the *S*s for whom it followed Fig. *J* ($P < 0.001$). Thus, the previous experience of Fig. *J* seems to have had about the same effect as its presence. In turn, Fig *J*, prevailingly flat in isolation, was described as cubical by more than half the *S*s for whom it followed Fig. *L* ($P < 0.001$). Thus, Fig. *L*, which exerted an insignificant influence across space, proved more effective across time. This finding points to an important peculiarity of temporally extended structures; namely, that later sub-units, upon arrival, interact with earlier ones as current percepts with memory-traces. Under circumstances which now require experimental definition, an item (such as Fig. *L*) may be more resistant to reorganization by neighboring items, and more prone to impose its organization on them instead, when it is present only as a trace than when it is visible.[15] Note, however, that such explorations of the differences between spatial and temporal structures become conceptually possible only after their similarity has been appreciated.

---

[14] This measure was taken to preclude possible effects of the current visible presence of the first description, although the results of a pilot experiment, in which the precaution was not taken, show that it made no measurable difference.

[15] Ongoing explorations of this feature of temporal structures have involved two procedures the results of which have some methodological relevance to the data reported above. The question can be asked whether, when successive figures are presented and each described before the next is shown, results indicating influence of an earlier on a later part of the sequence reflect the effect of one visual unit on another or merely the effect of the writing of one description on the writing of the next. When Experiment II*B* is repeated, with *S*s being asked to describe only the second of the two figures they have seen, the results are substantially the same as those of Table II, making it clear that the reported effect does not depend on the writing of descriptions of the earlier figure. Another question is whether an account in words perhaps reflects some purely verbal emphasis on a fleeting aspect of a complex visual percept rather than an expression of its stable and central character. When *S*s are asked to provide reproductions instead of descriptions, the original findings again are confirmed. Drawings of Fig. *J* which are two-dimensional and bilaterally symmetrical occur significantly more often, and three-

Summary

Two studies of phenomenal shape show that interaction between a spatial and a temporal structure can closely resemble interaction between two spatial structures. In the light of this resemblance, unique characteristics of temporal organizations invite examination. The main point to be made is that certain phenomena usually conceived of as effects of past experience on perception are subject to the same principles which govern perception itself.

---

dimensional drawings lacking bilateral symmetry occur significantly less often, when this figure has been preceded by Fig. *L* than when it has not been so preceded.

In a pilot version of this experiment, with the drawings larger in relation to card-size and to line-thickness, it was Fig. *J* (alone usually flat) which proved more labile (*i.e.* became usually cubic) when the two figures appeared together. Fig. *J* proved more influential across time (coming first) than across space (coming concurrently). Thus, it seems safe to say that, flat or cubical, the more labile of two figures profits more by temporal priority.

BACKGROUND REFLECTANCE AND THE CATEGORY-SCALING OF GRAY PAPERS

By Irwin Pollack, Cambridge, England

This note considers some stimulus- and procedural factors influencing the category-scaling of gray papers. The primary aim was to examine the role of the background-reflectance against which the papers were viewed, and study the interaction of background-reflectance with other factors. In particular, we wished to determine whether the entire categorical function for brightness is merely shifted, or whether there is a change in shape of the entire function, under different reflectances of background.[1]

Method

Stimulus-materials. The stimulus-materials were 29 gray papers, affixed to playing cards for rapid shuffling. The reflectances ranged from 72 to 3.5% as indicated in Table I.

TABLE I

Reflectances of Gray Papers Used in Scaling

Paper No.	Percentage reflectance	Paper No.	Percentage reflectance	Paper No.	Percentage reflectance
1[w]	72	11	31	21	10.1
2	66	12	30	22	9.2
3	64	13	30	23	7.4
4	58	14	24	24	5.7
5	58	15	21	25	5.7
6[i]	47	16	20	26	5.4
7	43	17	14	27	5.2
8	41	18	13	28	4.2
9	36	19	13	29[b]	3.5
10	32	20[i]	12		

[w] = white background employed in all tests.
[b] = black background employed in all tests.
[i] = intermediate backgrounds employed in tests in Fig. 1.

Viewing conditions. The stimulus-papers were placed upon background-papers 1, 6, 20 or 29, each 22 × 22 in. A 100-w. bulb, placed 25 in. over the materials

* Received for publication on June 19, 1962. The author was National Sciences Foundation Senior Postdoctoral Research Fellow at the Applied Psychology Research Institute. He acknowledges the assistance of Dr. Christopher Poulton in providing the gray papers and the calibrations.

[1] See W. C. Micheals and Harry Helson, A reformulation of the Fechner law in terms of adaptation-level applied to rating scale data, this Journal, 62, 1949, 355-368, especially Fig. 1, 263, and Table I, 362.

in a parabolic reflector, served as the main source of illumination. In tests employing filters, *O* observed the papers through a 1-in. hole in a black card over which filters were placed. In tests employing test-patches of different size, the linear dimensions varied from ⅛ that of a standard playing card (2.3 × 3.5 in.) to 4 times that of a playing card. In tests with an intermediate background, 0.16 × 0.22 in. test-patches were placed upon small black background, about 0.64 × 0.88 in., which, in turn, was placed upon the large white background.

The category-function for each testing condition was established separately. For example, to test the effect of background-reflectance, the category-function was determined for each reflectance of background on separate testing occasions. Objects were not viewed simultaneously against more than one background.

Observers. The *O*s were enlisted British Naval Ratings without previous experience in scaling experiments. Crews of 4 to 6 *O*s were available for two-week testing periods, usually 15 min. per day. All testing was carried out individually.

Category-Scaling Procedures

Category-naming. Single papers were placed one-at-a-time upon the background-sheet. The paper was removed after *O* responded. *O* was instructed to assign a number from '1' to '7' to each paper. Papers 1 (white) and 29 (black) were presented as examples of the lower end of Category 1 and the upper end of Category 7. In half of the tests represented in Fig. 1, the two papers representing the extreme reflectances were removed; in the other half of the tests, the extreme papers were not removed during category-scaling. Since there were no consistent differences between the results of the two conditions, the results were averaged. Subsets of 11 papers from the 29 papers were employed with a 7-point rating scale (unless specified otherwise).

In the examination of the effect of different background-reflectances, preliminary tests were carried out with the black and white backgrounds to insure a balanced distribution of papers over the entire brightness-scale for each background. From these results, balanced distributions of papers were derived for presentation against the intermediate backgrounds.

Category-sorting. A set of 11 papers was laid out upon the background-sheet before *O* in an ascending or descending order. *O* was instructed to arrange the cards in seven, or fewer, piles representing equal steps of brightness. Three distributions of 11 papers each were examined. Each set was weighted predominantly with either light, intermediate, or dark papers, respectively. The cumulative rank-order distributions of the papers are presented in the insert graphs of each of the top, middle, and bottom sections of Fig. 5. In an attempt to test the removal of the bias introduced by the weighted distributions, the method of iterative scaling was used in conjunction with the category-sorting tests. The method has been suggested by Stevens and associates.[3] It achieves a 'balanced' distribution of stimuli through a series of successive experiments. Specifically, the results of the first

[3] S. S. Stevens, On the averaging of data, *Science*, 121, 1955, 113-116; S. S. Stevens and E. H. Galanter, Ratio-scales and category scales for a dozen perceptual continua, *J. exp. Psychol.*, 54, 1957, 377-411; S. S. Stevens and E. C. Poulton, The estimation of loudness by unpracticed observers, *ibid.*, 51, 1956, 71-78.

category-scaling furnish the stimuli for the second scaling; the results of the second scaling furnish the stimuli for the third, etc. The purpose of the method is to achieve a category-scaling which is independent of the initial distribution of stimuli. Such a scaling is quickly achieved by the method of iterative scaling when the category-naming procedure is employed or when the category-sorting procedure is employed with stimulus-distributions sharing intermediate stimuli.[8]

Category-marking. The entire set of 29 papers was arranged upon the background-sheet either in an ascending or a descending order. *O* was provided n-1 intermediate markers, plus an end marker, with which to subdivide the 29 papers into n classes, approximately equally spaced in brightness. The procedure is formally identical with the category-sorting procedure except that the entire set of papers is viewed simultaneously.

Results

(A) Category-naming: (*1*) *Effect of background-reflectance.* Fig. 1 demonstrates that changes in the background-reflectance (parameter) alter the shape of the entire function of the brightness-category.

(2) Effect of brightness-level. Fig. 2 demonstrates that there is little change in the mean category-rating (ordinate) assigned to selected papers (parameter) with large changes in the absolute illumination-conditions (abscissa). The effect of background-reflectance (left vs. right sections) is maintained under a 1000:1 reduction in brightness-level. The only apparent change is the slight contraction of the scale when the papers are viewed against the white background under the extreme reductions imposed by the darkest filter.

(*3*) *Effect of test-area.* Fig. 3 demonstrates that there is little change in the mean category rating (ordinate) assigned to selected papers (parameter) with large changes in the area of the test patch (abscissa). The effect of background-reflectance (left vs. right sections) is maintained over a 1000:1 variation in area.

(*4*) *Effect of intermediate background.* The effect of adding a small intermediate black patch, interposed between the test-patch and the large white background, was to change the mean category rating about 0.4 rating units (except for the end stimuli). This change may be compared with an average change of about 1.0 rating unit between the white and black backgrounds in Fig. 1. Presumably, the intermediate black patch provided an effective contrasting border to the test-patch.

(*5*) *Effect of the number of rating categories.* Fig. 4 presents category-rating-functions with 5, 11, and 29 response-categories to the entire set of 29 papers. In addition, the lower right figure presents a magnified view of the middle range of the category-functions. It is noted that changes in the number of response-categories produce relatively small shifts in the shape of the brightness-category-function and in the effect of the white and black backgrounds.

---

[8] Irwin Pollack, Neutralization of stimulus bias in auditory rating scales, *J. Acoust. Soc. Amer.*, 36, 1964, 1272-1276; Iterative techniques for unbiased rating scales, *Quart. J. exp. Psychol.* (in press); Neutralization of stimulus bias in the rating of grays, *J. exp. Psychol.* (in press).

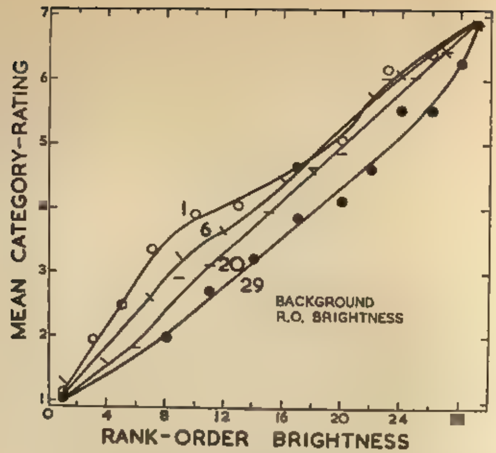

FIG. 1. EFFECT OF BACKGROUND-REFLECTANCE UPON THE BRIGHTNESS-CATEGORY SCALE FOR GRAY PAPERS

(The ordinate is the mean category-rating; the abscissa is the rank-order brightness of the gray papers; the parameter is background-reflectance, expressed in terms of the same rank-ordering. Each point represents two ratings for each of 11 *Os*.)

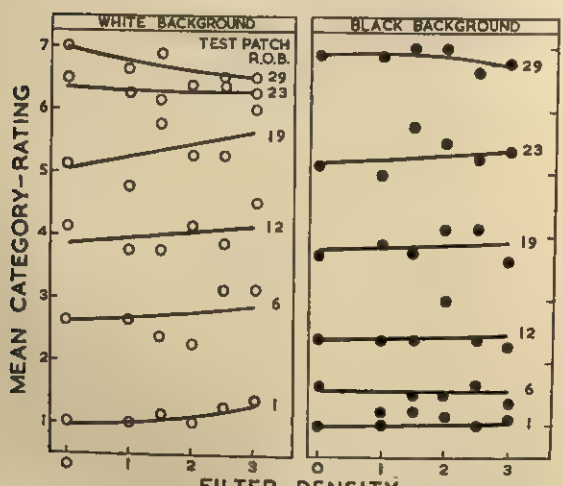

FIG. 2. EFFECT OF VIEWING GRAY PAPERS THROUGH A FILTER UPON THE SCALE OF BRIGHTNESS-CATEGORY

(The abscissa is filter-density. The parameter is the rank-order of brightness. The left section represents tests upon a white background; the right section represents tests upon a black background. Each point represents one rating by the category-naming procedure for each of 8 *Os*.)

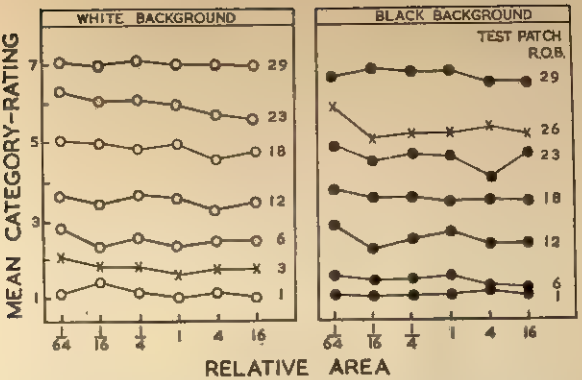

FIG. 3. EFFECT OF AREA OF TEST-PATCH UPON THE CATEGORY-SCALE BRIGHTNESS

(The abscissa is the relative area of the test-patch. One represents the size of a standard playing card. Each point represents one rating by each of 8 *O*s.)

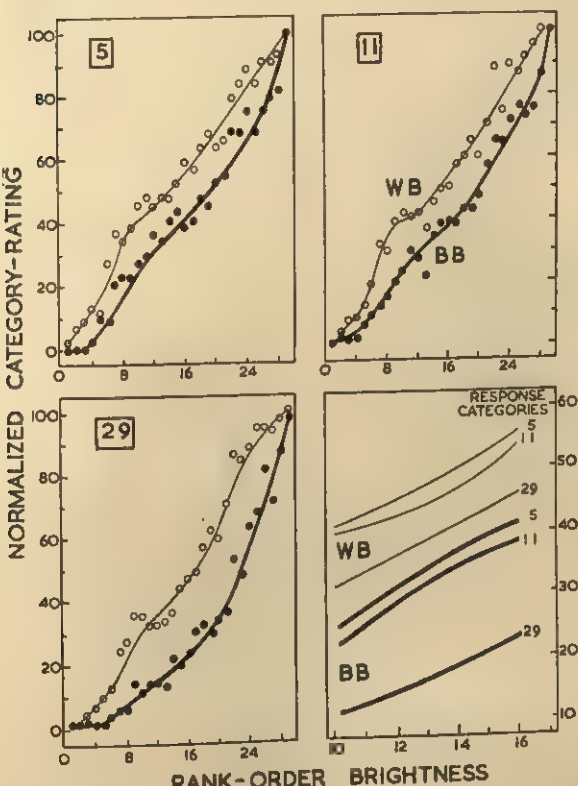

FIG. 4. EFFECT OF NUMBER OF RESPONSE-CATEGORIES ON THE BRIGHTNESS-CATEGORY FUNCTION

(The number of response-categories is given in the insert-box. The fourth section (lower right) expands portions of the separate functions. The ordinate has been normalized to accommodate 5, 7, 11, 17 or 29 categories. Each point represents one rating by each of 11 *O*s.)

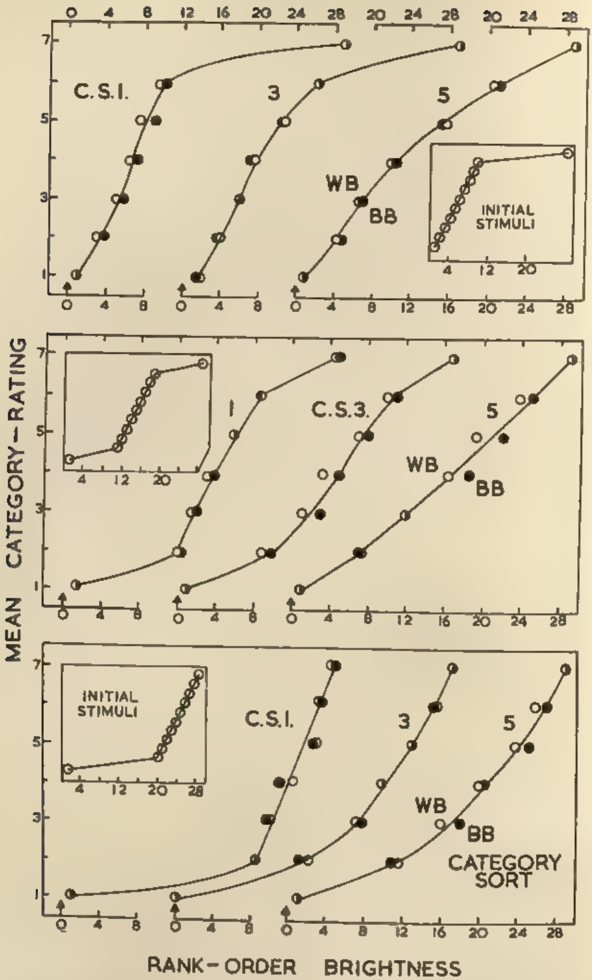

FIG. 5. EFFECT OF INITIALLY WEIGHTED STIMULUS-DISTRIBUTIONS UPON THE SCALE OF THE BRIGHTNESS-CATEGORY

(Initial stimulus-distribution shown in the insert. Successive curves represent results of the 1st, 3rd and 5th category-sortings with iterative scaling. Each point represents two category sortings by each of 11 *Os*.)

(B) Category-sorting. Fig. 5 presents the results of category-sorting with three initially weighted stimulus-distributions (insert graphs), employing the method of iterative scaling. Strong restrictions are imposed by the procedures in category-sorting. These restrictions are demonstrated by the fact that there is little difference between the ratings with white and

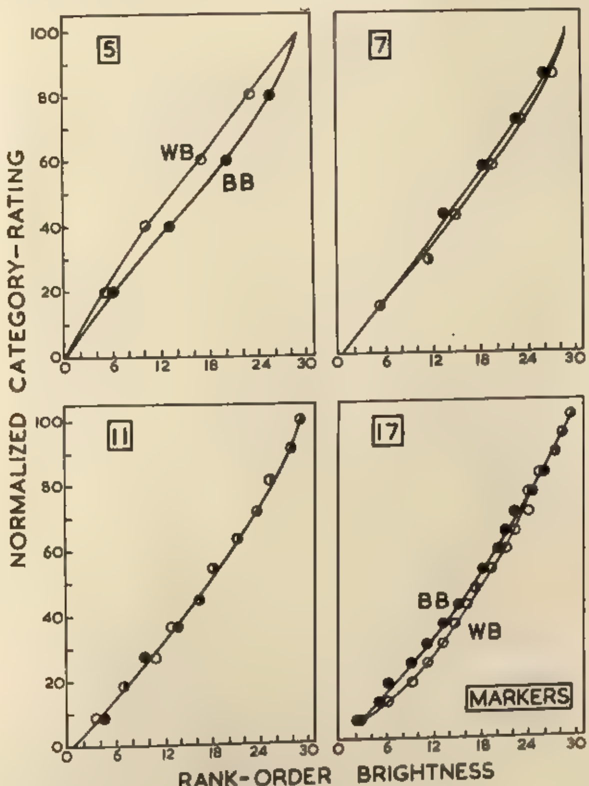

FIG. 6. EFFECT OF NUMBER OF MARKING-CATEGORIES ON THE BRIGHTNESS-CATEGORY SCALE

(The number of marking-categories is shown in the insert-box. Each point represents one marker-assignment by each of 6 *Os*.)

black backgrounds. Furthermore, even after five applications of the method of iterative scaling, the three category-functions derived from the three different initial distributions are still different. It may be noted that, within category-naming, differences among the same initial distributions

disappear within two successive tests of category-naming when the method of iterative scaling is applied.[4]

(C) Procedure of marker-scaling. The results of assigning markers to subdivide the set of 29 papers is presented in Fig. 6. Only with 5 marker-categories, *i.e.* with 4 intermediate markers, plus an 'end marker,' is there a consistent difference in the expected direction as a function of background.

Conclusions

Three conclusions, relating background-reflectance to the brightness-category scale, appear to be warranted from the present experiment:

(1) The effect of background-reflectance alters the shape of the category of brightness, rather than introducing a local deformity in the region of the background or introducing a shift of the entire function of the category.

(2) The effect of background-reflectance upon the scale of brightness is maintained over large variations in the area of the test-patch, the viewing illumination, and the number of response categories.

(3) The effect of background-reflectance is obscured by rating-scale procedures which impose severe restrictions upon the category-responses, *e.g.* the methods and procedures of category-sorting and marker-scaling.

[4] Pollack, Convergence upon stimulus bias in psychophysical scaling procedures (in preparation).

THE EFFECT OF COMMITMENT TO CORRECT AND INCORRECT DECISIONS ON CONFIDENCE IN A SEQUENTIAL DECISION-TASK

By NATHAN BRODY, Princeton University

Irwin and his associates have suggested that the classical theory of statistical inference may serve as a model of human decision-making.[1] The classical model of statistical inference makes no allowances for the *a priori* expectations of which of two alternative decisions is correct. On the other hand, the human consumer of the theory of statistical inference is frequently committed to one of the two alternative decisions. Pruitt has shown that commitment to a decision tends to increase the amount of information required to make a final decision.[2] The present study seeks to discover the effects of commitment to correct and incorrect initial decisions on subjective confidence in a sequential decision-making task.

METHOD

Subjects. Seventy-two Princeton High School boys (mean age 16.1 yr.) were used as *S*s in this experiment. They were recruited by the high school counseling service and were paid for participation in the experiment.

Procedure. The *S*s were randomly assigned to one of three experimental groups. Group D+ (initial decision correct), Group D— (initial decision incorrect), and Group ND (no decision) were the three groups. All *S*s were run individually. *E* started the session by telling *S* that he would be asked to take a test of ability to make decisions. *E* then read the following instructions for the ND group:

I am going to say the words plus and check 30 times. The words do not occur in any special order. One of these two words will occur 18 times, the other word will occur 12 times. I want you to make a decision about which of these two words will occur more often, *i.e.* which will be the word that occurs 18 times and which will be the word that occurs 12 times. You will be allowed only one

* Received for publication March 5, 1962. This research was supported by a grant from the National Institute of Health, Grant M-5697(A).

[1] F. W. Irwin and W. A. S. Smith, Further tests of theories of decision in an 'expanded judgment' situation, *J. exp. Psychol.*, 52, 1956, 345-348; Value, cost and information as determiners of decision, *ibid.*, 54, 1957, 229-232; F. W. Irwin, W. A. S. Smith, and J. F. Mayfield, Tests of two theories of decision in an 'expanded judgment' situation, *ibid.*, 51, 1956, 261-268. The issue of the relationship between these studies and the decision making processes of the scientist is discussed in the above articles.

[2] D. G. Pruitt, Informational requirements in making decisions, this JOURNAL, 74, 1961, 433-439; An exploratory study of individual differences in sequential decision making. Doctoral dissertation, Yale University, 1957.

decision. You can make this decision any time you want. Your performance will be evaluated by a scoring system. You will receive more points for a correct decision the earlier it is made. So, as you delay your decision you will have a better chance of getting the right answer. On the other hand, the longer you delay your decision, the fewer the number of points you will receive for the correct decision. Be sure to let me know when you are ready to make your decision.

The instructions were similar for Groups D+ and D—. Groups D+ and D— were asked, however, to make a decision before receiving any information. The final decision they made then contradicted or affirmed their original decision. The experiment was so rigged that the initial decisions of Group D+ were always correct and the initial decisions of Group D— were always incorrect.

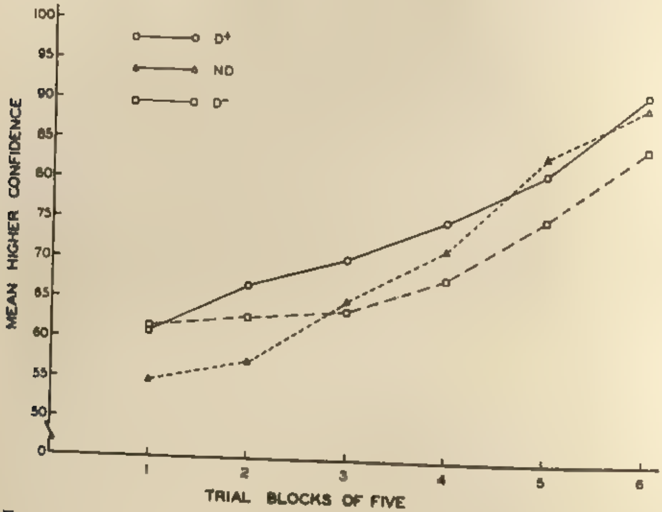

Fig. 1. Mean Higher Confidence over Trial-Blocks of Five

The stimulus-series used was the same random order of more frequent and less frequent stimuli for all the *S*s. Half of the *S*s in Group ND had the check stimulus occur 18 times. The stimulus selected to occur more often for Groups D+ and D— depended on the initial decision. *S* was supplied with pencil and paper and told that he could keep track of the stimuli presented to him.

After *S* had heard the first stimulus, *E* then asked *S* to state how confident he was in each of the two alternative decisions. The confidence-judgments were expressed in numerical units from 0 to 100, with the restriction that the sum of the two confidence-statements had to equal 100. Confidence-judgments were made after each stimulus-presentation.

Results

The higher of the two confidence-statements was recorded for each trial. Fig. 1 presents the mean of the higher confidence-statements by trial-blocks of five for each of the three experimental groups.

An inspection of Fig. 1 indicates that Groups D+ and D− tend to start the task with higher confidence than Group ND. Groups D+ and ND increase confidence at about the same rate, while Group D− tends to be slower in its rate of increase in confidence. Note further that Group D− never reaches the higher levels of confidence attained by the other two groups.

An analysis of variance was performed on the data presented in Fig. 1. The Trials variable was significant ($F = 87.19$ for 5 and 345 *df*, $p < 0.01$) and the Group by Trials interaction was significant ($F = 2.35$ for 10 and 345 *df*, $p < 0.01$). The Group by Trials interaction was partitioned into the linear, graduate, cubic, quartic, and quintic components of variance. The only component of trend that was significant was the linear component ($F = 8.85$ for 2 and 345 *df*, $p < 0.01$). The significant slope-interaction is largely attributable to the failure of

TABLE I

NUMBER OF *S*S ABOVE AND BELOW THE MEDIAN NUMBER OF TRIALS REQUIRED TO REACH 100% CONFIDENCE IN THE THREE EXPERIMENTAL CONDITIONS

	Above median	Below median
D+	8	16
ND	10	14
D−	17	7

$\chi^2 = 7.44$ for 2 *df*, $p < 0.05$.

Group D− to increase as rapidly in confidence as the other groups. The results of the analysis of the data presented in Fig. 1 are further supported by an analysis of the number of trials required to reach 100% confidence or certainty. Since all *Ss* did not reach 100% confidence, a X^2 median test is reported on these data. The data summarized in Table I, indicate that the *Ss* of Group D− are slower to attain 100% confidence.

The analysis presented in Fig. 1 is based on confidence statements over the entire course of the experiment. Consequently, it includes both pre- and post-decisions of confidence. It is of some interest to study changes in confidence that occur up to and including the trial of decision. Vincentized curves were constructed to study these data in the following manner. The trial of decision was noted for each *S*. The higher of the two confidence-statements was obtained at successive fifths of the number of trials to the trial of decision. Linear interpolation was used. The mean of the higher confidence at each fifth was obtained for each of the experimental groups. Fig. 2 presents these data.

Note again the tendency for the *Ss* of Groups D+ and ND to increase more rapidly in confidence than those of Group D−. An analysis of variance was performed on the data presented in Fig. 2. The Trials variable was significant ($F = 64.00$ for 4 and

276 *df*, $p < 0.01$). The Group by Trials interaction was also significant ($F = 3.02$ for 8 and 276 *df*, $p < 0.01$). The Group by Trials variability was again partitioned into the several components of trends by use of orthogonal polynomials. The linear component of trend was the only significant component ($F = 7.31$ for 2 and 276 *df*, $p < 0.01$). The analysis of variance of the data presented in Fig. 2 indicate that the significant slope-difference is not dependent on the failure of the *Ss* of Group

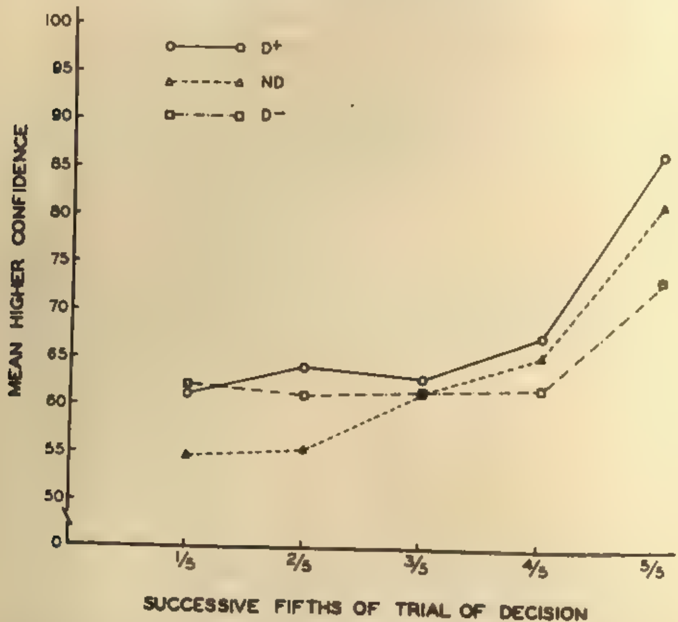

FIG. 2. MEAN HIGHER CONFIDENCE AT SUCCESSIVE FIFTHS OF THE TRIAL OF DECISION

D− to increase confidence rapidly after making their decisions. Rather, it indicates that they fail to increase confidence as rapidly prior to their decisions.

Table II presents mean values of the higher confidence at the initial trial, the higher confidence at the trial of decision, and the trial at which decision occurred for the three groups. In initial confidence only a comparison between the *Ss* of Groups D+ and D−, on the one hand, and the ND *Ss*, on the other hand, is significant, ($t = 2.02$ for 69 *df*, $p < 0.05$). This indicates that commitment to a decision tends to increase initial confidence.

In contrast, the analysis of the higher confidence at the trial of decision indicates that only the difference between the means of Groups D+ and D− is significant. Therefore, commitment to a correct initial decision leads to higher confidence at the trial of final decision than commitment to an incorrect initial decision.

There is no difference in the mean trial at which decision occurs among the three experimental groups. This result fails to duplicate the earlier finding of Pruitt that

the amount of information required to change a decision is greater than the amount of information required to make an initial decision.[3]

The final analysis reported is based on some suggestions by Pruitt about the parameters that are related to the timing of decisions in sequential decision-tasks.[4] He suggested that the rate of increase in confidence tends to be inversely related to the trial at which decision occurs, and the confidence-level at which *S* is willing to make a decision is positively related to the trial at which decision occurs. To obtain an estimate of the

TABLE II

MEANS FOR THREE EXPERIMENTAL GROUPS OF CONFIDENCE AT TRIAL ONE, CONFIDENCE AT TRIAL OF DECISION, AND TRIAL OF DECISION

Groups of *S*s	Confidence at Trial 1	Confidence at Trial of decision	Trial of decision
D+	62.3	86.5	19.3
ND	55.5	81.1	19.7
D−	60.9	73.5	20.2

TABLE III

CORRELATIONS AMONG RATE OF DECREASE IN CONFIDENCE. CONFIDENCE AT TRIAL OF DECISION, AND TRIAL OF DECISION

Correlation	Value	p
r_{12}	−.36	<.01
r_{23}	.23	<.05
r_{13}	.52	<.01
$r_{13.2}$	−.58	<.01
$r_{23.1}$	.52	<.01

1 = rate of increase in confidence.
2 = confidence at trial of decision.
3 = trial of decision.

rate of increase in confidence, the higher confidence-statements from the first trial to the trial preceding the trial of decision were noted for each *S*, and the slope of each *S*'s confidence-statements was obtained, using a least-square criterion. Table III reports the correlations among the trial of decision, confidence at the trial of decision, and the slope of the changes in confidence.

The data presented support Pruitt's suggestions. Rapid rate of increase in confidence tends to lead to early decision and high confidence at the

[3] Pruitt, *op. cit.*, 1961, 433-439. Pruitt used an explicitly formulated scoring system. The failure to replicate his results may be dependent on the specific number of points obtained for changing the initial decision.

[4] Pruitt, *op. cit.*, 1957.

trial of decision tends to delay the decision. Note also that the two parameters are not independent. Individuals who tend to increase confidence rapidly tend to have higher confidence at the trial of decision.

Discussion

It is tempting to treat the present experiment as an analogue to the process of decision-making in science. Although the comparison may not be legitimate, the resulting hypotheses about the process of scientific decision-making are interesting and may have heuristic value. The results would suggest the following hypotheses: A scientist committed to an incorrect hypothesis will require about the same amount of information to change his commitment as a scientist committed to the correct hypothesis. The scientist committeed to the incorrect hypothesis will, however, be slower to increase his confidence in the correct alternative on the basis of new evidence; will be less confident at the time he commits himself to the new hypothesis; and will take longer to reach 100% confidence or certainty in the correct hypothesis than the scientist initially committed to to the correct hypothesis.

The results of this experiment are interesting from a second point of view. They indicate that the commitment variable is a critical one. The simple act of stating one's expectations about which of two alternative decisions will be correct tends to subsequently influence confidence in that decision.

Summary

The present experiment explored the effects of initial commitment to correct and incorrect decisions on confidence in a task involving sequential decisions. Seventy-two high school boys were randomly assigned to one of three experimental groups: Group D−, the members of which made an initial decision which was invariably incorrect; Group D+, which made an initial decision that was invariably correct, and Group ND, which did not make an initial decision. The results support the following conclusions:

(1) Initial commitment to a decision tends to increase initial confidence.

(2) Initial commitment to an incorrect decision tends to decrease the rate of increase in confidence.

(3) Initial commitment to an incorrect decision leads to lower confidence in the final decision that initial commitment to a correct decision.

(4) Initial commitment to correct or incorrect decisions does not effect the timing of a final decision.

A STUDY OF THE VERBAL-TRANSFORMATION EFFECT

By THOMAS NATSOULAS, University of Wisconsin

The verbal-transformation effect has been defined by Warren as the occurrence of "illusory changes of distinct speech upon repetition."[1] A recording of an utterance is repeated at a constant rate again and again in identical form. The form being presented loudly and clearly gives way in *O*'s perception to new forms, often very different from the original. During a few minutes a number of forms may be heard, involving many transformations. Warren and Gregory seem to have discovered the phenomenon.[2] Warren studied the effects of masking noise, loudness, rate of presentation, length, and complexity of the repeated utterance,[3] and he has compared aged with young adults as well as monaural and binaural conditions of listening.[4] Axelrod and Thompson correlated the number of transformations with reversals of perspective.[5] Taylor and Henning found marked effects on the kind of forms heard, and on the number of transformations, of instructions to the *O*s that they would hear only English words.[6] The present experiment compares meaningful and meaningless bisyllables and introduces a trial-by-trial procedure permitting observations to be made of the effect of an interval during which repetitions cease.

METHOD

Materials. A Wollensack (T-1500) tape-recorder set at extreme treble was used to present repetitive auditory stimulation through earphones to *O*. Nine bisyllabic

* Received for publication January 6, 1964 this research was supported in part by the Research Committee of the Graduate School of the University of Wisconsin with funds provided by the Wisconsin Alumni Research Foundation. I thank Dr. Willard R. Thurlow for the use of his equipment and his help in estimating the loudness of the stimuli, and Eleanor Ryvkin, John T. Murphy, and John E. Ross for help with the statistical analyses.

[1] R. M. Warren, Illusory changes of distinct speech upon repetition—The verbal transformation effect, *Brit. J. Psychol.*, 52, 1961, 249-258.

[2] R. M. Warren and R. L. Gregory, An auditory analogue of the visual reversible figure, this JOURNAL, 71, 1958, 612-613.

[3] Warren, *op. cit.*, 249-258.

[4] Warren, Illusory changes in repeated words: Differences between young adults and the aged, this JOURNAL, 74, 1961, 506-516; An example of more accurate auditory perception in the aged, in Clark Tibbitts and Wilma Donahue (eds.), *Social and Psychological Aspects of Aging*, 1962, 789-794.

[5] Seymour Axelrod and Larry Thompson, On visual changes of reversible figures and auditory changes in meaning, this JOURNAL, 75, 1962, 673-674.

[6] M. M. Taylor and G. B. Henning, Verbal transformations and an effect of instructional bias on perception, *Canad. J. Psychol.*, 17, 1963, 210-223.

utterances were the auditory stimuli, each of them presented in the form of the identical utterance repeated at the rate of 65 times per min. To produce this effect, a short piece of tape with a single utterance recorded on it was cut and spliced noiselessly into a loop, and then as the loop played on and on, the utterance was repeatedly recorded onto a long reel of tape by means of another tape-recorder of the same type. Throughout the experiment, the stimulus-materials were presented binaurally at a monaural loudness of 68 phons.[7]

Four of the nine bisyllables were common English words: *parrot, dollar, seven,* and *bottom.* Four more, nonsense-words included as controls for the English words, were pronounced in a way as similar to the words as possible; to produce the control-bisyllables the first and last letters of each English word were transposed in position. They are referred to henceforth as *nc-parrot, nc-dollar, nc-seven,* and *nc-bottom,* where *nc-* is an abbreviation for 'nonsense-control for.' The ninth bisyllable, used on practice-trials only, was *nc-nigger.*

Observers. Thirty-two men and an equal number of women served as *Os.*

Procedure. Upon arrival, *O* was given a copy of the following instructions, which were read aloud by *E.*

You will hear a voice through the earphones repeating the same word over and over. Listen carefully to the voice. If you hear a change to another word, signal by tapping your pencil on the desk, as soon as the change occurs. If there is a change back to the original word or to yet another word, signal again as soon as it occurs. *Any change* to a new word or to a word previously heard should be signalled right away. Even if the change is momentary, *i.e.* involving just one repetition, be sure to signal it. Do not wait to confirm the change by listening further. Do not worry about whether what you hear is meaningful or familiar.

The word which is being repeated will *not* actually change. The same word will be repeated again and again. Whether changes are heard or not heard when none in fact occurs on the tape is the question of this study.

The procedure is divided into a series of trials. At the end of each trial, the tape will be stopped, giving you a short time to write as many of the words you just heard as you can remember. Please use one page from this pad for each trial.

Do not worry about being right or wrong! I am not interested in how well you can identify what is on the tape. What is important is to find out what you are hearing under these special conditions. There is *nothing preferable* about hearing changes or not hearing them. This study is not designed to evaluate you but to gather knowledge with your help. Just respond according to what you experience.

Listen carefully but passively. Do not *try* to produce or not produce changes in the word you are hearing.

O's questions were answered and his attention was called again to the second paragraph of the instructions. A practice-word, *nc-nigger* for all *O*'s, was presented for five trials of 60 repetitions each. After the first trial, during which *O* had tapped his pencil, or after the third trial, if he had not tapped as yet, instructions were given to the effect that *E* was not concerned with changes of accent or emphasis from one part of the word to another, that tapping should occur only if a change in

[7] The latter value was determined by matching the loudness of a tone of 1000 c.p.s. presented to one ear with the loudness of an utterance presented to the other, and by finding the threshold for hearing the tone with the same ear as in the matching one.

sound occurred. This was defined for *O* as "a change in one or more sounds of the word, the addition of one or more sounds, the dropping out of one or more sounds." Following the fourth practice-trial, further instructions were given about signalling. It was pointed out that a transformation may not persist; it may be present for only one repetition and be replaced on the very next repetition. *O* was asked to be sure to signal both times, when the change to a different form occurred, and then again when there was a change back or to something new. It was stressed that highly transient transformations may occur but need not occur for many people.

O then proceeded through the experiment proper which consisted of five trials of 60 repetitions each for each of four bisyllables, with 15 sec. between trials and 1–2 min. between bisyllables. At the end of the experimental session, *O* was asked to pronounce aloud all the forms he had written after each trial. *E* corrected the spelling or clarified the writing wherever necessary. Where differences in spelling reflected no differences in pronunciation, the spelling was corrected to be the same.

Design. Half the men and half the women were randomly assigned to an experimental group which heard the four bisyllables comprising Word-List I: *dollar, parrot, nc-bottom,* and *nc-seven.* The remaining *O*'s heard the items of Word-List II: *seven, bottom, nc-dollar,* and *nc-parrot.* Each of these four groups of *Os* was divided further into subgroups of four *Os* each, the subgroups differing with respect to the order in which the bisyllables were presented to them. The four sequences of bisyllables were selected on the basis of a 4×4 Latin square, the same square being used for Word-List I and Word-List II by categorizing *seven* and *nc-seven,* for example, as the same sound-pattern. In addition there are in the analyses to be reported the within-*O* effects of Position (of sound-pattern, first, second, third, or fourth), Sound-Patterns (*seven* or *nc-seven* vs. *bottom* or *nc-bottom* vs. *dollar* or *nc-dollar* vs. *parrot* or *nc-parrot*), Trials, and Thirds (of a trial). The factor of Sex is omitted from the reported analyses because it has no significant effects on any of the measures used either as a main effect or in interaction with other factors.

Results

Satiation-period. Howard has introduced the concept of satiation-period in the context of perspective-reversals of rotating wire-figures.[8] For present purposes, the satiation-period was defined as the number of repetitions of an utterance prior to the first transformation reported.[9] The scores for satiation-period used in statistical analysis were log-transformed to reduce skewness. A 'mixed' analysis of variance was performed with repeated measures on each *O* according to a Latin square as described in the section

[8] I. P. Howard, An investigation of a satiation process in the reversible perspective of revolving skeletal shapes, *Quart. J. exp. Psychol.*, 13, 1961, 19-33.

[9] In the case of the 14 *Os* who reported no transformations for 1 (7 *Os*), 2 (3 *Os*), or 3 (4 *Os*) bisyllables, a satiation-period of 300 repetitions was assigned. The instances of hearing no change over 300 repetitions were distributed according to bisyllables as follows: *bottom*, 2; *nc-bottom*, 5; *dollar*, 6; *nc-dollar*, 2; *parrot*, 4; *nc-parrot*, 1; *seven*, 4; *nc-seven*, 1.

on design. Three statistically significant effects emerged from the analysis of variance of log satiation-period: Sound-Patterns at $p < 0.001$ ($F = 8.13$, $df = 3$ and 168), Position at $p < 0.01$ ($F = 3.80$, $df = 3$ and 168), and Word-Lists × Sound-Patterns at $p < 0.005$ ($F = 4.58$, $df = 3$ and 168). *Parrot* and *nc-parrot* appear to be the most readily transformed (mean log satiation-period 1.511) ; *seven* and *nc-seven,* and *bottom* and *nc-bottom* are intermediate (1.652 and 1.666 respectively); and *dollar* and *nc-dollar* require the greatest mean log satiation-period (1.810).

The significant main effect of Position indicates that ths position of a sound-pattern, whether it is the first, second, third, or fourth one to which *O* is exposed, affects his satiation-period; the mean log satiation-period declines from 1.754 to 1.663 to 1.550, but there is a lengthening of the period for the fourth bisyllable, (1.672).

The interaction of Word-Lists × Sound-Patterns was expected to be significant on the hypothesis that meaningfulness of a sound-pattern will prolong its satiation-period relative to the nonsense-control for that sound-pattern. There were two bases for this expectation: (a) A meaningful word has been encountered repeatedly in everyday life and recognized under a variety of conditions, a variety of pronunciations, and with a variety of background-noise conditions. Thus a wider latitude of variation in the sound of the word should be acceptable to *O* as the 'same' word than in the case of the respective nonsense-control. (b) The auto-inhibition' produced by repeated stimulation must overcome a positive tendency to hear the word veridically. This positive (excitatory) tendency is a function of frequency of exposure to the word in *O*'s past experience.

Table I presents in its first two columns the mean log satiation-period for each word and its nonsense-control. It can be seen that in three cases the meaningful word has a longer satiation-period than its control. These differences are significant by Duncan's test in the case of *dollar* vs. *nc-dollar* ($p < 0.05$) and *parrot* vs. *nc-parrot* ($p < 0.001$). It was not expected originally that the effect would hold for all cases of a word compared to its nonsense-control. This transposition does not produce the very same phonetics as in the original.

Forms. An analysis of variance identical to that using log satiation-periods as scores was performed on the total number of different forms heard over the five trials for each bisyllable. Both the main effect of Sound-Patterns and the interaction of Word-Lists × Sound-Patterns again were highly significant: Sound-Patterns at $p < 0.001$ ($F = 6.81$, $df = 3$ and

168) and Word-Lists × Sount-Patterns also at less than $p < 0.001$ ($F = 9.27$, $df = 3$ and 168).

The right half of Table I shows the mean number of different forms heard over five trials for each of the eight bisyllables used in the experiment. In each case, the meaningful word yields fewer different forms than the nonsense-controls. Duncan's test indicates that *parrot* and *dollar* differ from their nonsense-controls at $p < 0.001$ and $p < 0.005$, respectively. These results are consistent with expectations based on reasoning similar to that for satiation-period: More different forms of the nonsense-controls are reported because less of a transformation is necessary for the new form

TABLE I

MEANS FOR MEANINGFUL BISYLLABLES AND NONSENSE-CONTROLS

	Satiation-period (log repetitions)			Number of different forms		
	meaningful	control	both	meaningful	control	both
bottom	1.693	1.640	1.666	3.8	4.3	4.1
dollar	1.908	1.712	1.810	3.3	4.6	3.9
parrot	1.691	1.330	1.511	4.3	5.7	5.0
seven	1.638	1.666	1.652	3.7	4.4	4.1

to be considered different from one heard earlier. Furthermore, meaningful words may be more stable, allowing less opportunity for new forms to appear; this is true at least for the satiation-period.

Number of transformations. An analysis of variance, parallel with one exception to those performed on the log satiation-periods and the number of forms, was made for a score based on the number of verbal transformations signalled. The exception refers to the addition of two, within-*O* factors—Trials and Thirds of a trial. In the first third of a trial, 19 transformations were possible, in the second and third, 20. For every *O*, the number of verbal transformations for every third was expressed as a proportion of the number possible. These proportions were converted by arc-sine to angles, and the analysis of variance was based on the transformed scores.

Both Trials and Thirds were significant main effects at $p < 0.001$ (respectively, $F = 53.85$ with $df = 4$ and 224, and $F = 118.12$ with $df = 2$ and 112). The corresponding means expressed in angles for Trials were 10.697, 16.443, 19.911, 21.929, 22.317. Within trials, the number of transformations increased similarly: 12.819, 20.242, 21.719.

The results for Trials and Thirds are in keeping with a conception of the underlying mechanism for the verbal-transformation effect which empha-

sizes the growth of an auto-inhibitory process. As repetitions proceed, the original form and subsequent forms become unstable and are superceded by new forms or by forms heard earlier. The latter recur by virtue of a recovery from satiation relative to the remaining forms. The rate of trans-

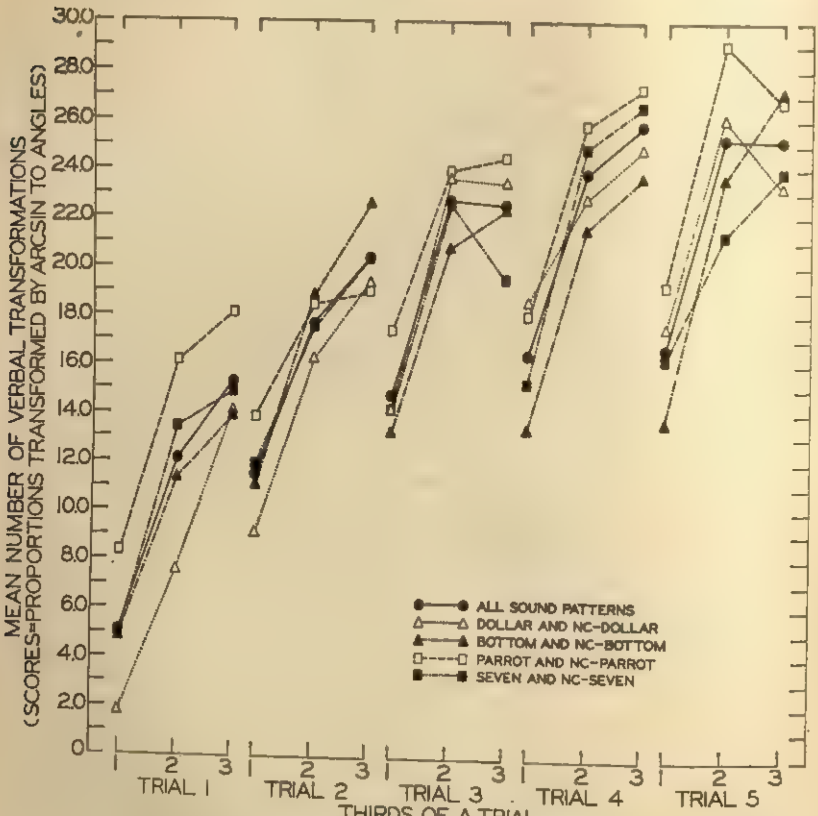

FIG. 1. THE MEAN PROPORTION OF REPORTED VERBAL TRANSPORTATIONS RELATIVE TO THE NUMBER POSSIBLE (TRANSFORMED TO ANGLES BY ARC-SINE) FOR THE THIRDS OF EACH TRIAL

formation increases with repetitions as a result of an increase in the number of forms on the balancing edge of excitation-inhibition.

As may be seen in Fig. 1, and as indicated by the insignificant interaction of Trials × Thirds, the pattern described in terms of the accumulation of satiation repeats itself on all trials and for all sound-patterns with few exceptions. A marginally significant interaction of Trials × Thirds ×

Sound-Pattern ($F = 1.62$, $df = 24$ and 768; $p < 0.05$) is attributable to the few instances in which there is a drop from the second to the third thirds.

Recovery from auto-inhibition is evidenced by the effect of the 15-sec. delay between trials. Fig. 1 shows that the mean number of transformations reported during the first third of a trial is less for all sound-patterns than the mean number reported during the final third of the previous trial. A Duncan multiple-range test performed on the means for all sound-patterns (solid line) indicates a significant decline in number of transformations following each delay ($p < 0.05$, < 0.005, < 0.001 and < 0.001 for the four successive delays, respectively).

The main effect of Sound-Patterns is statistically significant ($F = 2.69$, $df = 3$ and 168, $p < 0.05$), as is the interaction of Word-Lists × Sound-Patterns × Trials ($F = 2.93$, $df = 12$ and 672, $p < 0.001$). *Parrot* and *nc-parrot* is the sound-pattern for which there is the greatest mean number of transformations, while the remaining three sound-patterns are practically equal. The significant three-way interaction does not lend itself to an interpretation in terms of meaningfullness.

The latter results taken with the fact of an insignificant interaction of Word-Lists × Sound-Patterns strongly suggests, in the light of the results on satiation-period, that once the satiation-period is over (once the initial form has reached an unstable state), its meaningfulness is no longer a determinant of the number of verbal-transformations. This is a consequence of the hypothesized state of affairs at the point of the first transformation: the excitatory process representing the original form is balanced by an auto-inhibitory process at that point. Whatever the habitual strength of the excitatory process, the result is the same once it is overcome by inhibition. From then on, at least for a time, the frequency of changes will depend on the strengths of the excitatory processes of the forms serving to replace the original one.

TEMPORAL CHROMATIC INDUCTION: THE INTERACTION OF TWO SUCCESSIVE PULSES MEASURED BY SUBJECTIVE ESTIMATES

By ROBERT KESTON, New York State Psychiatric Institute

The purpose of this investigation was to determine the changes in perceived color produced by the temporal succession of two brief, superimposed pulses of light (ranging in duration from 10-250 m.sec.). The measurement of perceived color was limited to subjective reports of hue and saturation.

The sensation elicited by a flash of light is complex and often persists beyond the cessation of the physical stimulus. The primary response to a colored flash of light is frequently followed by a series of after-images. This series usually consists of an alternation of positive and complementary images.[1]

A particularly interesting after-image situation was studied by Bidwell in 1896.[2] Various colored materials were viewed through the open sector of a slowly rotating disk. In addition to the open sector, the disk had a white sector and a black sector. Under appropriate conditions, *only* the complementary after-image of the colored material was perceived. For instance, if a red ribbon was viewed through the open sector, the ribbon appeared not red at all, but only in the green after-image color. It has since been found that this after-image is seen only at relatively high levels of illumination. As the over-all luminance of the disk is reduced, the after image 'disappears,' and the positive image is perceived.[3]

Sperling re-investigated the Bidwell after-image phenomenon, using a double-flash technique involving the repeated cycling of a colored flash (*ca.* 20 m.sec.), a 'white' flash (*ca.* 60 m.sec.), and a dark interval (*ca.* 80

* Received for publication January 13, 1964 this research was supported in part by National Science Foundation Grant G-9594. The author is indebted to Dr. Leo M. Hurvich, Dorothea Jameson, and Dr. J. Rutschmann for advice and assistance.

[1] D. B. Judd, A quantitative investigation of the Purkinje after-image, this JOURNAL, 38, 1927, 507.

[2] Shelford Bidwell, On subjective color phenomena attending sudden changes of illumination, *Proc. Roy. Soc. Lond.*, 60, 1896-7, 368-377. For an historical review see M. E. Erb and K. M. Dallenbach, Subjective colors from line-patterns, this JOURNAL, 52, 1939, 227-241.

[3] H. E. Lehmann, Preliminary report on a device for the objective measurement of the negative after-image phenomenon, *Science*, 112, 1950, 199-201.

m.sec.).[4] He was able to produce complementary after-images of various stimulus-configurations "without prior positive image." Sperling also investigated the range of lumiances and durations over which the phenomenon takes place. The following stimulus-configuration was recycled every 500 m.sec.: A 'white' masking flash, subtending $6 \times 10°$ visual angle (v.a.), was presented for 75 m.sec. The onset of this flash was defined as time = 0.0. At some time during the cycle, a test-flash (duration less than 0.10 m.sec.) was superimposed on the masking field. "The energy of this test flash and its time of occurrence relative to the onset of the masking field were varied systematically."[5] A single *O* reported (1) whether he had seen the test-flash; (2) whether the test-flash (if seen) was brighter or darker than the masking field; and (3) whether he saw a complementary after-image. The results indicated that the complementary after-image was seen only when the test-flash preceded or coincided with the masking field in time, and only when its luminance was considerably higher (2.5 to 3.5 log units) than that of the masking field. Absolute luminance-levels were not reported.

EXPERIMENT I: PULSE-DURATION

Since repetitive cycling phenomena of this kind are very complex and difficult to analyze, single presentations of paired flashes were investigated in the present study. A phenomenon similar to that investigated by Bidwell does occur, in that the complementary after-image of the first flash often is perceived. The phenomenon is systematically dependent on the durations of the paired flashes and on the interval between them.

Method. The stimulus consisted of a red flash followed by a near-white flash. Both pulses were presented to the same area of the retina, and they were always of equal luminance and duration. There was no dark interval between the pulses.

Sylvania glow-modulator tubes, operated at 20 ma. rms, served as stimulus-sources. The first pulse was transmitted through a Schott interference-filter with maximal transmission at 645 mμ and a half-width of 12 mμ (according to manufacturer's specifications). The second was filtered only by neutral-density filters.

The durations of the pulses were controlled by electronic R-C timers. The electrical signals produced by these timers were calibrated with a Hewlett-Packard interval meter and printer. The durations varied by no more than 3% from the values stated on the dials, with repeat-reliability of approximately ±50 μsec. In addition, it was found that the light-pulses produced by the glow-modualtor tubes accurately followed the electrical signals produced by the timers. The rise- and fall-times are less than 250 μsec.

---

[4] George Sperling, Negative afterimage without prior positive image, *Science*, 131, 1960, 1613-1614.
[5] Sperling, *op. cit.*, 1613.

The visual field is illustrated at the top of Fig. 1 (*A*). It consisted of a 3° v.a. foveal test-patch surrounded by a dark annulus and a near-white surround. The luminance of the surround was fixed at 23 mL, and it was present throughout the experiment.

Three luminances of the paired pulses—21.0, 2.1 and 0.21 mL—and eleven pulse-durations—ranging from 5–150 m.sec.—were investigated. The luminances of the paired flashes were equated by means of flicker photometry. The actual lumi-

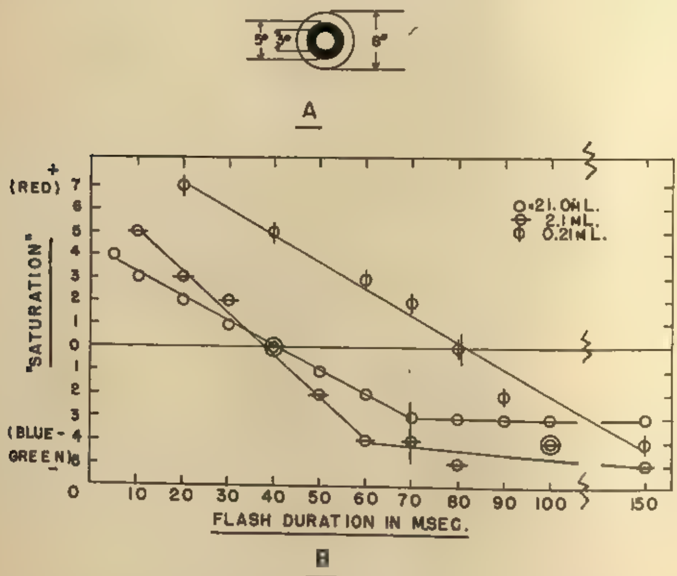

FIG. 1. (*A*) VISUAL FIELD. (*B*) RELATION BETWEEN SATURATION AND DURATION

(Luminance is the parameter. The major vertical bars indicate the first pulse durations at which two flashes were reported.)

nance-levels were determined by heterochromatic matches to a white, diffuse surface, which was subsequently calibrated with the Macbeth Illuminometer.

Five *Os* reported the number of flashes perceived and were asked to rate the 'saturation' of the second flash or of the single flash when the two flashes were fused into a single impression.[6] A standard flash-pair that appeared red and fused was arbitrarily given a value of +5. White was represented by zero, and negative values were assigned to blue-greens (complementary hue of the red). Thus, a subjective scale was established on which positive values represent red hues of increasing saturation, zero (0) represents white, and negative values represent blue-

[6] The *Os* were three women and two men, ranging in age from 20-36 yr., with normal color vision according to the American Optical-Hardy-Rand-Rittler color-test. All *Os* were trained for at least 3 hr. in 'saturation' scaling.

greens of increasing saturation. Although the validity of this scale as a quantitative measure of saturation *per se* may be questioned, it was found to yield consistent results.

An experimental series consisted of the presentation of all pulse-durations in ascending or descending order at one luminance-level. Each session consisted of six series, one ascending and one descending for each luminance-level. Each *O* participated in four experimental sessions.

Results. The results are presented graphically in Fig. 1 (*B*). 'Saturation' is plotted on the ordinate and the duration of the first pulse on the abscissa. Luminance is the parameter. Each point denotes the median of 40 judgments representing five *O*s. The lines were fitted by inspection. The vertical bars indicate the first duration at which *O*s report two flashes.

Clearly, an effect similar to the Bidwell phenomenon does occur, but it is part of a continuum: As the pulse durations vary from 1-150 m.sec., the 'saturation' rating varies from + 4 or + 5 to 0, and finally to − 3 or − 4. It should be noted that in the case of the 21.0 and 2.1 mL functions, there is a region between 40 and 70 m.sec. in which a fused impression is reported, but in which that impression is blue-green in hue. Thus, in this region no sensation of red is reported.

The results for 21.0 and 2.1 mL are similar. At 0.21 mL, however, the pulse-durations at which *O* first reports blue-green, as well as the pulse-durations at which the flashes are discriminated, are greater. Furthermore, there is no region in which the flashes are both fused and appear blue-green.

Discussion. According to the Hurvich-Jameson development of the Hering color theory, color vision is subserved by three opponent-processes.[7] These authors state that "the two members of each pair are opponent, both in terms of the opposite nature of the assumed physiological processes and in terms of the mutually exclusive sensory qualities."[8] Two of these processes are chromatic (red-green and yellow-blue) and one is achromatic (white-black). Stimulation of one pole of an opponent process, for instance red, results in decreased responsiveness of that pole and increased responsiveness of the opponent pole, in this case green. There is a tendency toward an equilibrium-state associated with a sensation of neutral grey.

Mutual interaction or induction among various elements of the visual field are conceived to be important determinants of the visual response. Hurvich and Jameson have predicted and demonstrated that, in spatial interactions affecting perceived color, induced responses are both opponent

---

[7] L. M. Hurvich and Dorothea Jameson, An opponent-process theory of color vision. *Psychol. Rev.*, 64, 1957, 384-404.

[8] Hurvich and Jameson, *op cit.*, 385.

to and proportional to the magnitude of the inducing stimulus.[9] That is, yellow or red stimulation in a surrounding area will result in the induction of blue or green activity in the focal area, and conversely. These authors have suggested that analogous effects occur in the temporal realm.[10]

The results of Experiment I bear out this prediction and are consistent with the above findings with respect to spatial interactions: If induced opponent responses are proportional to the magnitude of the inducing stimulus, in this case red, and this magnitude increases with duration, then at short durations the induced activity is small, becomes increasingly greater balancing the primary red-activity at the white point, and at longer durations the green activity seems to predominate. This is, of course, an oversimplified generalization, and considerable experimental work will be required before the parallel between spatial and temporal induction-effects can be stated with any precision.

Experiment II: Dark Interval

In this experiment, the effect of varying the dark interval between the pulses was investigated.

Method. The luminances of the two pulses were fixed at 21.0 mL, and 10–16 dark intervals, varying in duration from 1–250 m.sec., were explored. Three durations of the paired pulses—5, 40, and 100 m.sec.—were investigated. An experimental series consisted of the presentation of all dark intervals in ascending or descending order at one pulse-duration. Each session consisted of three series, one for each pulse-duration. The dark interval was varied an equal number of times in ascending and descending order. Each *O* took part in six experimental sessions. In all other respects the method was the same as that employed in Experiment I.

Results. The results are presented in Fig. 2. 'Saturation' is plotted on the ordinate and the dark interval on the abscissa. Pulse-duration is the parameter. Each point denotes the median of 30 judgments representing 5 *O*s.

At the 5 m.sec. pulse-duration, as the dark interval varies from 1-250 m.sec., the 'saturation' rating varies from +5 through 0 to −3 or −4 and, finally, again to 0. A descending branch (1-100 m.sec.) and an ascending branch (100-250 m.sec.) of the function can be discriminated. Thus, under these conditions, variation of the dark interval produces a transition in perceived color (red to blue-green) similar to that brought about by variation of pulse-duration in Experiment I. The function relating 'saturation' to dark interval is apparently U-shaped at the 5-m.sec. pulse-duration,

[9] Jameson and Hurvich, Opponent chromatic induction: Experimental evaluation and theoretical account, *J. opt. Soc. Amer.*, 51, 1961, 46-53.

[10] Jameson and Hurvich, Perceived color and its dependence on focal, surrounding, and preceding stimulus variables, *J. opt. Soc. Amer.*, 49, 1959, 890-898.

and, as the pulse-duration increases (40 and 100 m.sec.), the entire function appears to be shifted to the left. The initial point, dark interval = 0, is defined by the results of Experiment I.

The interaction between pulse-duration and dark interval is illustrated

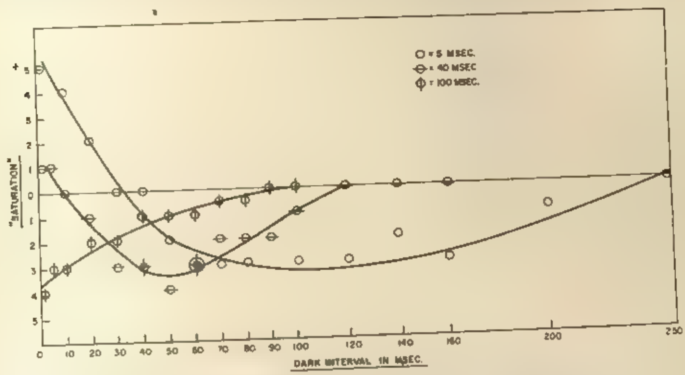

FIG. 2. RELATION BETWEEN SATURATION AND DARK INTERVAL
(The parameter is pulse duration. Vertical bars indicate the first dark intervals at which two flashes were reported.)

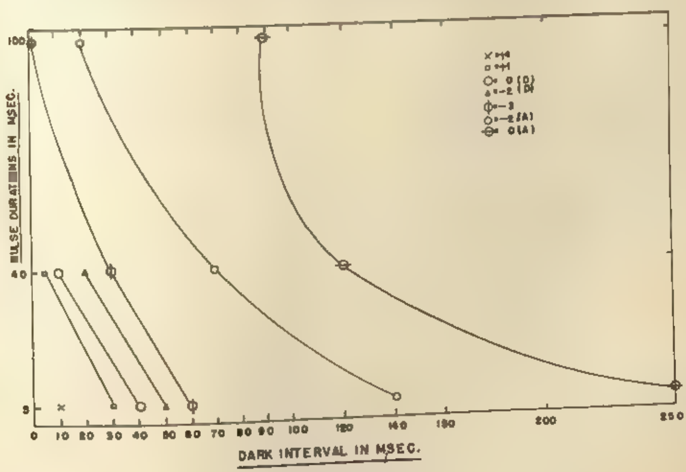

FIG. 3. ISO-SATURATION CONTOURS: PULSE-DURATION VS. DARK INTERVAL
(Saturation is the parameter.)

by Fig. 3, in which the former is plotted as ordinate and the latter as abscissa. Here 'saturation' is the parameter. These curves may be thought of as 'iso-saturation' contours. Apparently, the dark interval required for a particular (constant) 'saturation' *decreases* as the pulse-duration *increases*.

In an investigation of spatial visual interactions, Jameson and Hurvich found that the magnitude of the induced responses (the amount of change in the focal stimulus) was proportional to the degree of continguity between the focal area and the inducing areas.[11] In other words, an area of 'darkness' interposed between the focal area and the inducing area served to reduce the interaction. At first glimpse it seems reasonable to suppose that the dark interval in a temporal-interaction situation might serve the same function, but the present study clearly demonstrates that this is not necessarily so. The dark interval may serve to enhance the degree of interaction, to decrease the interaction, or it may have no effect, depending on various other conditions; that is, dark interval interacts with various other parameters in determining perceived color in the temporal-interaction situation.

Summary

In Experiment I, the interaction of two successive flashes of equal luminance and duration was studied as a function of pulse-duration and luminance of the paired flashes. The first flash alone was yellowish-red in appearance and the second near-white. The results indicated that as the paired pulse-durations lengthen, reds become increasingly desaturated, then white, and finally blue-green. These results were shown to be consistent with Hurvich and Jameson's finding that, in spatial interactions affecting perceived color, induced responses are both opponent to and proportional to the magnitude of the inducing stimulus.

In Experiment II, the effect of introducing a dark interval between the flashes was investigated. It was found that variation of the dark interval could produce a transition in perceived color similar to that brought about by variation of pulse-duration in Experiment I. In addition, the interaction between dark interval and pulse-duration was specified.

[11] Jameson and Hurvich, Opponent chromatic induction: Experimental evaluation and theoretical account, *J. opt. Soc. Amer.*, 51, 1961, 46-53.

TORQUE: A NEW DIMENSION IN TACTILE-KINESTHETIC SENSITIVITY

By BURTON WOODRUFF and HARRY HELSON,
Kansas State University

Recent studies have shown that the effective stimulus in lifted-weights is a function of the density as well as the weight of the stimulus. Nyssen and Bourdon, using the method of limits, found that lighter, denser stimuli were equal to heavier, less dense stimuli,[1] and Furth, using the comparative rating-scale, showed that point of subjective equality (*PSE*) varied with the density of the standard when the weights of the series-stimuli were held constant.[2] In this study, a third factor influencing judged weight is investigated, viz. torque. Although this factor enters into operations of twisting and turning knobs, dials, handwheels, and all manipulanda in which the applied force is tangential to the line of motion, sensitivity to torque has never, to our knowledge, been determined. Everyday experience shows that a rod held by one end is harder to support in a horizontal than in a vertical position. Although its weight is the same in both positions, there is an added force, torque, defined as the turning moment of a tangential effort, which is produced when the rod is held in the horizontal position.

Sensitivity to torque was measured in this study by using a stimulus of constant weight which was manipulated to vary its torque according to its position on a rod. Results of two experiments utilizing the method of constant stimulus-differences and the comparative rating scale are reported.

EXPERIMENT I

Method: Subjects. In the first experiment, 10 men (2 undergraduate and 8 graduate students in psychology) were employed as *Ss*. Their age-range was from 20 to 28 yr. The *Ss* had no prior information about the nature of this study other than the fact it involved 'lifted weights' in the traditional sense.

---

* Received for publication May 22, 1964. Preparation of this paper was supported in part by the Office of Naval Research under Contract Nonr-3634 (01) with Kansas State University.

[1] R. Nyssen and J. Bourdon, A new contribution to the experimental study of the size-weight illusion, *Acta Psychologica*, 12, 1956, 157-173.

[2] H. G. Furth, The effect of the size-weight illusion on adaptation-level, *J. exp. Psychol.*, 60, 1960, 150-154. Cf. also L. B. Hoisington, On the non-visual perception of the length of lifted rods, this JOURNAL, 31, 1920, 114-146.

Apparatus. The nature of the problem made it possible to use one easily adjustable stimulus. The stimulus consisted of a round wooden rod 1.3 cm. in diameter, 28.1 cm. long, and having a mass of 27.5 gm. The second part of the stimulus was a wooden block, 4 × 5 × 7.25 cm., hung from notches in the rod by a 20-gauge 5-in. copper wire. The wire was looped over the top of the block along the 5-cm. axis. Nails were driven into the bottom of the block to make it weigh 121.7 gm.; hence, the total mass of the stimulus was 149.2 gm. A 1.125-in. diameter fiber washer was glued on the rod 2.2 cm. from one end. The first notch in the rod was 8.8 cm. from the washer and beyond this were four other notches 2.5 cm. apart. By hanging the block on the rod at each of the five notches, five

TABLE I

VALUE OF STIMULUS (TORQUE) AS A FUNCTION OF POSITION OF THE MOVABLE WEIGHT

	Position of weight from fulcrum				
	9.7 cm.	12.2 cm.	14.7 cm.	17.2 cm.	19.7 cm.
Total torque (gm.-cm.)	1538.0	1842.2	2146.5	2450.7	2754.9

stimuli differing in torque were obtained. The torque exerted on *S*'s fingers when the block was placed at each of the five positions is given in Table I.

The stimulus was enclosed in a cardboard-box, 10 × 18 × 11.5 in., to obstruct *S*'s view of the stimulus-conditions. The back of the box was so cut away that *E* could adjust the stimulus, which protruded from the front of the box through a centered slot. The slot was 1.25-in. wide and 6.75-in. long (vertical axis), and its bottom was 5.5 in. from the bottom of the box. When not in use, the washer end of the rod rested on the bottom of the slot and the other end rested on a 5.5-in. high cardboard-partition glued 8 in. from the front of the masking box. All of the apparatus was painted flat black.

Procedure. The *S*s were seated at a table of convenient height. The masking box and the stimulus were set in front of the *S*s in a manner that prevented them from seeing directly into the slot. They were cautioned against looking into the box and were not aware of the conditions of stimulation.

The stimuli were presented according to the method of paired comparisons, in which the torque of each stimulus was compared with the torque of every other stimulus. The presentation-order was standard (*Std*) first, variable (*Co*) second. The *S*s made their judgments with reference to the comparative rating-scale suggested by Helson, Michels, and Sturgeon,[3] utilizing the following categories with the numbers in parentheses assigned for purposes of computation: very very much heavier (10), very much heavier (9), much heavier (8), heavier (7), little heavier (6), equal (5), little lighter (4), lighter (3), much lighter (2), very much lighter (1), and very very much lighter (0). The categories, without numbers, were in view of *S*s for reference, if desired.

The *S*s were instructed to lift the stimuli by putting the first joint of their index fingers under the rod parallel and as close to the washer as possible, and their

[3] Harry Helson, W. C. Michels, and Artie Sturgeon, The use of comparative rating scales for the evaluation of psychophysical data, this JOURNAL, 67, 1954, 321-326.

thumbs on top of and parallel to the rod. They were to lift and lower the stimulus freely without hesitation. Judgments were made after the second stimulus was lifted.

They were given 10 practice trials to acquaint them with the range of stimulation, after which they were given 75 trials. The positions of the block on the rod were randomized within blocks of five trials using a given position as *Std.* The data are averages of three series of five blocks each. The order of standards and comparisons within blocks and series was randomized differently for each *S* to minimize order-effects. There was a 1- or 2-min. rest-period at the end of each series.

Stimuli. The rod on which the stimuli were hung can be considered as a simple lever with its fulcrum 0.3 cm. behind the washer. Observation showed that this was approximately where the weight of the rod rested on the index finger. Since the fiber washer was so close to the fulcrum and the rod extended slightly in back

TABLE II

EXPERIMENTAL AND CALCULATED JUDGMENTS FOR EACH STIMULUS AND THE ADAPTATION-LEVELS (*ALs*) WITH EACH STANDARD

Standard (gm-cm.)	Stimulus-series (gm.-cm.) 1538		1842		2146		2451		2755		*A*	*b*
	Obs.	Calc.*	Obs.	Calc.*	Obs.	Calc.*	Obs.	Calc.*	Obs.	Calc.*		
1538	5.0	5.0	6.2	6.4	7.3	7.4	8.4	8.2	8.8	8.9	1545	− .615
1842	3.6	3.5	5.1	5.2	6.3	6.5	7.8	7.5	8.2	8.3	1795	− .858
2146	2.8	2.5	4.0	4.3	5.6	5.7	6.9	6.8	7.6	7.6	1980	− .798
2451	1.8	2.1	3.4	3.2	4.2	4.4	5.0	5.4	6.4	6.2	2319	− .477
2755	1.7	1.9	2.6	2.5	3.7	3.3	4.4	4.3	5.2	5.5	2631	+3.191

* The calculated values in this table were obtained from the following equation: $J_{calc.}=10(X-A)/[X+(1+b)/(1-b)A]+5$, where *X* refers to series-stimuli, *A* is the adaptation-level, and *b* is the *y*-intercept

of the fulcrum, these two items probably canceled. They were, therefore, omitted in the analysis of the stimulus.

The rod can be considered as a lever weighing 27.7 gm. and 26-cm. long with a fulcrum at one end. Physically, the rod acts as if it were a massless lever, 13 cm. long with a mass of 27.5 gm. hung on the very end. The torque produced by this lever is the product of the mass (27.5 gm.) and the distance over which the mass acts (13 cm.). Thus, the rod, by itself, produces a torque of 357.7 gm-cm. To hold this rod level, *S* must counteract this torque by an equivalent force.

The total torque of the system is the sum of that produced by the rod itself and the stimulus-block. When the 121.7-gm. stimulus is hung on the notch closest to the washer (first stimulus-position), it adds a torque of 1180.5 gm.-cm. (9.7 cm. × 121.7 gm.); hence to hold the rod level with the stimulus in the first position, *S* must exert a force of 1538.0 gm.-cm. Table I gives the torque for each of the five stimulus-positions which determined the value of both the *Std* and *Co* stimuli.[4]

Results and discussion. The observed and fitted mean judgments and the adaptation-levels (*ALs*) for the five *Std* are given in Table II. The data

[4] We are indebted to Professor P. L. Miller of Kansas State University for the analysis and evaluation of the stimulus-conditions.

in Table II have been fitted to a negatively accelerated function proposed by Helson which yields *ALs* when absolute or comparative-rating categories are used.[5] Furth obtained both *AL* and *PSE* (by the usual methods) and found that the two measures were practically identical.[6] The data for the highest *Std* (2755 gm.-cm.), in which the judgments were predominately below 5, were fitted by reversing the numerical scale, *i.e.* the observed

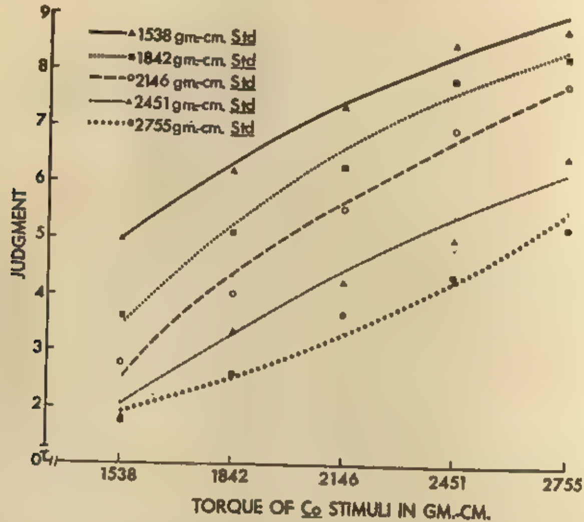

FIG. 1. JUDGMENTS OF TORQUE AND BEST-FITTING CURVES WITH LOW, MEDIUM, AND HIGH VALUES OF STANDARDS

values were subtracted from the topmost value of the scale, 10, thus giving a positively accelerated function. The observed data and curves of best fit are shown in Fig. 1.

The curves in Fig. 1 show that sensitivity to torque yields functions similar to those obtained when only tactile pressure is involved. With standards at the low end of the stimulus-series, negatively accelerated curves fit the data obtained and, as the standard moves toward the high end of the series, the curves become less negatively accelerated until, with the highest *Std* (2755 gm.-cm.), the curve is positively accelerated. These

[5] Harry Helson, Adaptation-level as a basis for a quantitative theory of frames of reference, *Psychol. Rev.*, 55, 1948, 297-313. For examples illustrating fitting this function and for the additional information it yields, see Harry Helson and Philip Himelstein, A short method for calculating the adaptation-level for absolute and comparative rating methods, this JOURNAL, 68, 1955, 631-637.

[6] Furth, *op. cit.*, 150-154.

findings parallel Guilford's in which lifted-weights were employed and each member of the series served as standard.[7] The small discrepancies between experimental and calculated values in Table II attest the excellence of the fit of the function chosen to represent the data. In every case, except that involving the lowest standard, 1538 gm.-cm., *AL* has a value below that of the standard, as is usual with lifted weight data. The difference between the standard and *AL* in the case of the 1538-gm.-cm. *Std* (8 gm.-cm.) is relatively so small as to be insignificant. Taking *AL* as a measure of point of subjective equality, since it represents the stimulus judged equal to the standard, we find that torque, like lifted-weights, gives rise to negative time-order effect (*TOE*), on the whole.

Data obtained by the method of paired comparisons and comparative-rating scales can usually be converted to percentages for the purpose of calculating Urban thresholds. When the data for this study were converted to percentages, the resulting ogives were very steep. Evidently, the step-intervals between stimuli were considerably larger than one *DL;* so the decision was made to repeat this experiment, using smaller step-intervals and to employ the usual three-category judgments, heavier, equal and lighter, to obtain Urban limens.

Experiment II

Method: Subjects. Five naïve undergraduates (men) served as *Ss* in this experiment, and their ages were 18 or 19 yr.

Apparatus. Additional notches were so cut in the rod that the stimuli for this study were separated by half the interval used in the first study. The distance between stimuli was 1.25 cm. instead of 2.50 cm. making the difference between adjacent stimuli 152.125 gm.-cm. instead of 304.25 gm.-cm. All other aspects of the apparatus remained the same.

The middle value of the series-stimuli was still 2146 gm.-cm. but there were seven *Co* stimuli instead of four, with the middle stimulus as *Std.* The values of the *Co* in round numbers were 1690 gm.-cm., 1842 gm.-cm., 1994 gm.-cm., 2146 gm.-cm., 2299 gm.-cm., 2451 gm.-cm., and 2603 gm.-cm.

Procedure. The same procedure was used as previously, except that only three categories, heavier, equal and lighter, were used as in the method of constant-stimulus differences. The middle stimulus, 2146 gm.-cm. was the *Std* and the order of presentation was *Std* first, *Co* second.

One series of judgments was made by each *S* to allow him to become acquainted with the range of stimuli, after which 10 series of judgments were recorded as the experimental data. Seventy judgments were thus made by each *S* and with the five *Ss*, the percentages for each stimulus were based on 50 observations. The series-stimuli were randomized differently for each *S* to minimize order-effects.

[7] J. P. Guilford, *Psychometric Methods,* 1954, 158.

Results and discussion. The pooled responses of the five *Ss* were used to determine the constants of the best-fitting ogives for the lighter and heavier judgments by the Urban method to determine *DLs*, as shown in Fig. 2.[8] The 50% point for lighter yields a *DL* of 1942 gm.-cm. and that for heavier, a *DL* of 2241 gm.-cm. The Weber fraction of relative *DL* for lighter is thus 9.48% and for heavier, 4.42%. *PSE*, calculated as one-half the interval of uncertainty, is 2092 gm.-cm. This value of *PSE* is not far

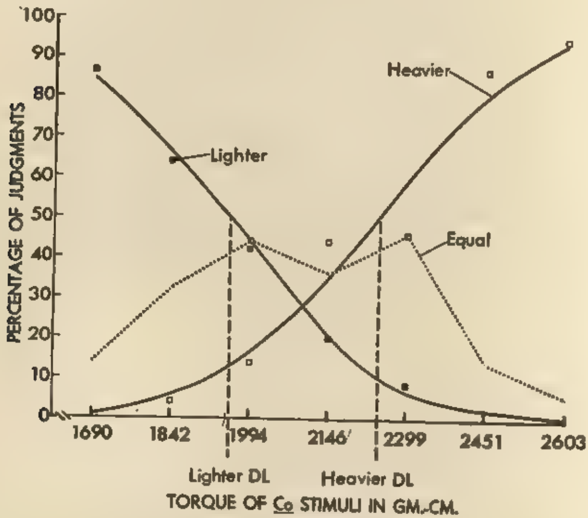

FIG. 2. UPPER AND LOWER *DLs* FOR TORQUE CALCULATED BY URBAN'S METHOD AND THE BEST-FITTING OGIVES WHEN 2146 GM.-CM. IS USED AS THE *Std.*

from the value found for *AL* when the middle stimulus was used as standard (1980 gm.-cm.) in the first experiment. When the differences in the spacing of the stimuli, the methods of judging, and the use of different functions for fitting the data in the two experiments are considered, the agreement between them is excellent in view of the fact that *PSE* and *AL* differ by only a little over 5%.

SUMMARY AND CONCLUSION

The two experiments reported here show that sensitivity to torque measured by two different psychophysical methods behaves in a manner similar to sensitivity to tactile pressure. Torque involves not only psychological

[8] Cf. Guilford, *op. cit.*, 129-131.

pressure but also resistance to movement, hence a greater kinesthetic component than is usually found in lifted-weight stimulation *per se*. Negatively accelerated curves for low values of the *Std* and positively accelerated curves found for high values of *Std* in lifted-weight data by Guilford were also found here when sensitivity to torque was measured by the comparative-rating scale. The Weber fraction for heavier judgments is close to that reported for weight-judgments, but that for lighter judgments was found to be considerably larger, perhaps because only data for one time-order of *Std* and *Co* were used in obtaining the Urban upper and lower thresholds. The values of both *AL* and *PSE* indicated that negative *TOE* is operative in judgments of torque as well as in lifted-weight judgments obtained by traditional methods. In conclusion, it should be pointed out that there was some arm- as well as finger- and wrist-movement in the experiments reported here and that quite different measures of sensitivity to torque may be obtained if only finger- and wrist-movement is involved, as is true when small knobs and dials are turned.

RESISTANCE TO EXTINCTION IN THE PIGEON AS A FUNCTION OF SECONDARY REINFORCEMENT AND PATTERN OF PARTIAL REINFORCEMENT

By R. C. GONZALEZ, VIRGIL GRAF, and M. E. BITTERMAN,
Bryn Mawr College

Suppose that rats are partially reinforced in a runway with differentiated end-boxes (*e.g.* black on reinforced trials, white on unreinforced trials) which cannot be seen until the instrumental response has been made, and then extinguished—one group with the end-box used on reinforced training trials, and a second group with the end-box used on unreinforced training trials. The principle of secondary reinforcement suggests that resistance to extinction will be greater in the first group than in the second; yet exactly the opposite results were obtained in two experiments reported about 10 years ago,[1] and the results of other experiments, while to some extent inconsistent, certainly do not yield readily to a simple secondary-reinforcement interpretation. Hulse and Stanley obtained fragmentary evidence of greater resistance in the group extinguished under end-box conditions like those of reinforced training trials, but the design of their experiment was unbalanced[2] Lawrence and Festinger,[3] like Freides,[4] and Notterman,[5] found no significant difference between the groups. Using two different negative end-boxes, Marx found no significant differences between groups extinguished with one or both of the negative end-boxes and a group extinguished with the posi-

* Received for publication April 2, 1964. This research was supported under terms of Contract Nonr 2829(01) with the office of Naval Research and by Grant MH-9577 from the National Institute of Mental Health.

[1] M. E. Bitterman, W. E. Fedderson, and D. W. Tyler, Secondary reinforcement and the discrimination hypothesis, this JOURNAL, 66, 1953, 456-464; C. B. Elam, D. W. Tyler, and M. E. Bitterman, A further study of secondary reinforcement and the discrimination hypothesis, *J. comp. physiol. Psychol.*, 47, 1954, 381-384.

[2] S. H. Hulse, Jr., and W. C. Stanley, Extinction by omission of food as related to partial and secondary reinforcement, *J. exp. Psychol.*, 52, 1956, 221-227.

[3] D. H. Lawrence and Leon Festinger, *Deterrents and Reinforcement*, 1962, 106-107.

[4] David Freides, Goal-box cues and pattern of reinforcement, *J. exp. Psychol.*, 53, 1957, 361-371.

[5] J. M. Notterman, A study of some relations among aperiodic reinforcement, discrimination training, and secondary reinforcement, *J. exp. Psychol.*, 41, 1951, 161-169.

tive end-box, but a group extinguished with all three end-boxes (the positive and the two negatives) was significantly more resistant than any of the others.[6] Most recently, Elmes has added a third experiment showing significantly greater resistance in the group extinguished with the negative end-box than in the group extinguished with the positive, although (contrary to Marx) a group extinguished with both the positive and the negative end-boxes was intermediate in resistance between the other two.[7] Whatever the explanation of the discrepancies, it seems reasonable to suppose that the effect of secondary reinforcement tends to be masked in such experiments by some resistance-producing factor associated with the presentation of the negative end-box.

We report here an experiment of the same general design, but with another animal (the pigeon) and another apparatus (with key-pecking instead of running as the instrumental response, and illumination of the grain-magazine by lamps of different color on reinforced and unreinforced trials substituted for the different end-boxes of the runway). The pigeon was not chosen primarily in the expectation that it might give different results than the rat (although some broadening of the phyletic base of the research certainly seemed desirable), but on methodological grounds. The pigeon-situation was easier to automate, and it afforded an interesting measure of consummatory response to the goal-stimuli which, by association with reinforcement or nonreinforcement, were presumed to acquire differential secondary-reinforcing properties. In the experiment by Freides, speed of running in the positive and negative end-boxes was measured and found to be different despite the fact that resistance to extinction of the antecedent runway-response was not significantly affected by the end-box which was used.[8] In the pigeon-situation, the goal-response measured was the insertion of the animal's head into the lighted aperture of the grain-magazine.

Another novel feature of the present experiment was that the pattern of reinforcement was varied in a factorial design: Gellermann's sequences of reinforced and unreinforced trials were used for half the animals,[9] while for the remaining animals *extended* sequences—which involved much longer runs of unreinforced trials—were used. On the basis of

---

[6] M. H. Marx, Resistance to extinction as a function of degree of reproduction of training conditions, *J. exp. Psychol.*, 59, 1960, 337-342.
[7] D. G. Elmes, The role of frustration in the extinction of a running response, *Psychon. Sci.*, 1, 1964, 345-346.
[8] Freides, *op. cit.*, 361-371.
[9] L. W. Gellermann, Chance orders of alternating stimuli in visual discrimination experiments, *J. genet. Psychol.*, 42, 1933, 206-208.

previous work with fish[10] and rat,[11] it was anticipated that the extended sequences would produce greater over-all resistance to extinction.[12] One purpose of the variation in pattern of reinforcement was to compare the effects of positive and negative magazine-colors under conditions making for high and low over-all resistance, the notion being that an interaction might help to explain some of the discrepancy in previous results. A second purpose was to look at the effects of long runs of unreinforced trials in a situation which permitted measurement also of the goal-response.

Method: Subjects. The *Ss* were 40 experimentally naïve White King pigeons obtained from a local supplier. After 10 days of ad-lib feeding, they were reduced over a period of three weeks to 75% of their satiated weights, and then their training was begun.

Procedure. The *Ss* were trained in a ventilated picnic chest containing a panel in which were mounted a Gerbrands grain-magazine and a single pigeon-key. The magazine and key were centered on the panel, the magazine just above floor-level, and the key 7 in. above the floor. The key was illuminated with white light; the magazine could be illuminated either with red or with green light. In the lower lip of the magazine-aperture was a photocell which was used to detect the insertion of *S*'s head into the aperture. Illumination of the photocell was provided by the magazine-light. After magazine-training (half the animals with the red light and half with the green light), the animals were trained to peck the key. After the pecking response was established, each animal was given three daily sessions consisting of 20 reinforced trials. Each trial began, after a 6-sec. intertrial-interval (*ITI*) in darkness, with the illumination of the key. A single peck turned off the key-light and initiated a 4-sec. magazine-cycle. The magazine-light (red for *Ss* magazine-trained with red light and green for *Ss* magazine-trained with green light) was turned on at the start of the cycle, but the grain-tray was not presented until the animal inserted its head into the magazine-aperture. At the end of the magazine-cycle, the next *ITI* began. Two time-measurements were made on each trial—the latency of the instrumental response (key-pecking) and the latency of the goal-response (insertion of the head into the magazine-aperture). These times were read by *E* from 0.01-sec. Standard Electric timers. (The other events of the experiment were programmed automatically.)

On the basis of their performance in this preliminary training, the *Ss* were divided into two matched groups—Gellermann and Extended—and given 20 addi-

[10] R. C. Gonzalez, R. M. Eskin, and M. E. Bitterman, Further experiments on partial reinforcement in the fish, this JOURNAL, 76, 1963, 366-375.

[11] Gonzalez and Bitterman, Resistance to extinction in the rat as a function of percentage and distribution of reinforcement, *J. comp. physiol. Psychol.*, 58, 1964, 258-263.

[12] A runway-experiment with the pigeon in which the intertrial interval was 24 hr. showed no effect of extension on resistance to extinction (W. A. Roberts, D. H. Bullock, and M. E. Bitterman, Resistance to extinction in the pigeon after partially reinforced instrumental training under discrete-trials conditions, this JOURNAL, 76, 1963, 353-365), but there was every reason to expect that the massed-trials result for fish and rat would be duplicated in the pigeon.

tional days of training (Days 4-23). There were 20 trials per day, of which 10 were reinforced exactly as in the last stage of the preliminary training. On the remaining 10 trials, which were not reinforced, response to the key produced a magazine-light of the opposite color (*e.g.* red if the color on positive trials was green), and the goal-response failed to produce the grain-tray. For the Gellermann animals, reinforced and unreinforced trials were scheduled in accordance with selected Gellermann orders. For the Extended animals, longer runs of unreinforced trials (as long as 14 in succession) were scheduled, with the restriction that the ratio of reward must be exactly 50% in each pair of days.

On the basis of their performance in training, the animals in each of the two groups were divided into two matched subgroups (*S* and *N*) and extinguished.

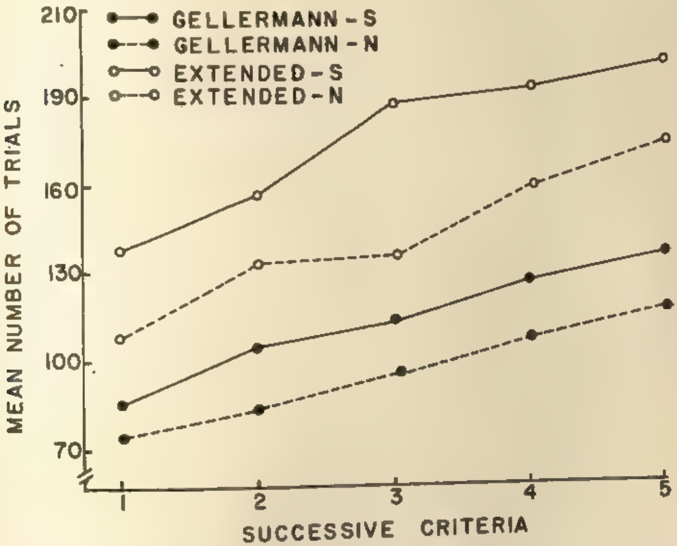

FIG. 1. EXTINCTION OF THE INSTRUMENTAL RESPONSE IN TERMS OF THE NUMBER OF TRIALS TO REACH CRITERIA OF INCREASING SEVERITY

During extinction, only one of the two magazine-lights was used for each animal. The color for animals in the *S*-subgroups was that associated with reinforcement in training, while the color for the *N*-subgroups was that associated with nonreinforcement in training. There were 20 trials per day to a criterion of five successive failures to make the instrumental response in 30 sec. When, on any trial, the instrumental response was not made in 30 sec., the trial terminated and the next *ITI* began. When the instrumental response was made, an extinction-trial was just like an unreinforced training trial except, in the case of the *S*-groups, for the color of the magazine-light.

Results. The extinction of the instrumental response is plotted in Fig. 1 in terms of mean trials to criteria of progressively increasing severity.

(1-5 successive failures to respond in 30 sec.). The four curves, which are essentially parallel, show both a pattern-effect (the Extended curves are higher than the Gellermann) and a secondary-reinforcement effect (the *S*-curves are above the *N*-curves). Analysis of variance confirms these impressions. For pattern of reinforcement, $F = 30.9$ with 1 and 36 *df*, $p < 0.01$; for secondary reinforcement, $F = 5.7$ with 1 and 36 *df*, $p < 0.05$. None of the interactions approaches significance at the 5% level.

With respect, then, to the effect of the pattern of partial reinforcement on resistance to extinction, the present results for the pigeon are like those for the rat. The failure to find such an effect in a previous experiment with the pigeon in which the intertrial interval was 24 hr. may have been due simply to insufficient training, although the possibility of an interaction between pattern of partial reinforcement and spacing of trials should not be overlooked.[18] With respect to secondary reinforcement, however, the present results are different from the modal results for the rat. Whether we are dealing here with a phyletic difference, or whether the contrasting results for the two animals are to be understood in terms of a difference in some situational variable of as-yet-

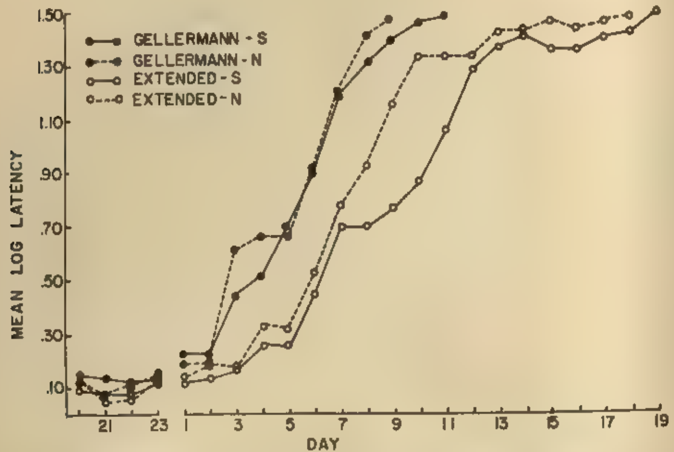

FIG. 2. MEAN LOG LATENCY OF THE INSTRUMENTAL RESPONSE IN THE FINAL STAGE OF ACQUISITION AND IN EXTINCTION

[18] Roberts, Bullock, and Bitterman, *op. cit.*, 363-364.

unsuspected importance, it is too early to say. What is clear only is that if an anti-secondary-reinforcement effect does operate in the pigeon, as it seems to in the rat, the present data provide no indication of it.

Daily mean log latencies in the final phase of acquisition and in extinction are plotted in Fig. 2 for the instrumental response and in Fig. 3 for the goal-response. The principal resemblance between the two sets of extinction-curves is that the N-latencies are greater than the S-latencies in each case. The effect is not very striking in the instrumental curves for the Gellermann groups, which, considered in relation to the criterion-data

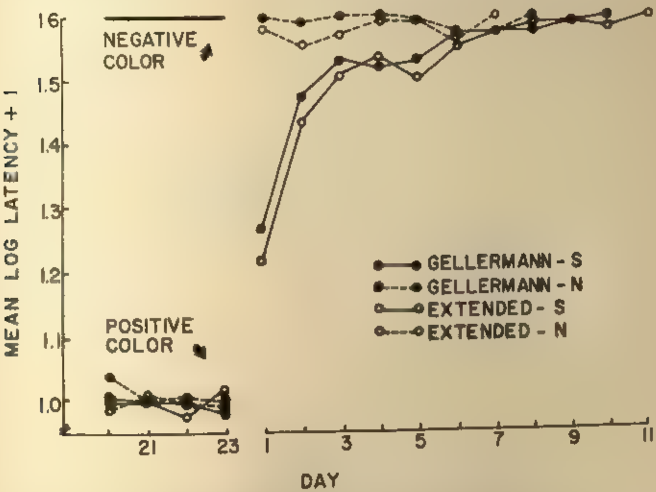

FIG. 3. MEAN LOG LATENCY OF THE GOAL-RESPONSE IN THE FINAL STAGE OF ACQUISITION AND IN EXTINCTION

of Fig. 1, suggests that secondary reinforcement served mainly (at least in the Gellermann case) to reduce the number of response-failures.

An impressive difference between the two sets of curves (for instrumental and goal-response) is in over-all rate of extinction. The goal-response-curves for the positive color show a considerable amount of extinction on the first day, while the corresponding instrumental curves show no appreciable extinction until the third (Gellermann) or fourth (Extended) day. Clearly, the instrumental responses-curves are diverging during a period in which the goal-response curves are converging. It seems clear, then, that the greater resistance to extinction of the instrumental response when it is followed by the positive as compared with

the negative stimulus is independent (at least to a large extent) of the actual occurrence of the goal-response to that stimulus during extinction. Another impressive difference between the two sets of curves is that the extended effect, which predominates in the instrumental curves, is negligible in the goal-curves, which show the effect of secondary reinforcement predominantly. The greater resistance to extinction of the instrumental response after extended as compared with Gellermann reinforcement apparently is independent of the frequency with which the goal-response occurs in extinction.

Summary

In a discrete-trials partial-reinforcement experiment of factorial design, the pigeon's resistance to extinction was found to be a function of secondary reinforcement (resistance was greater when the color of the magazine-light was that used on reinforced rather than on unreinforced training trials) and of the pattern of partial reinforcement (resistance was greater after long as compared with short runs of unreinforced trials in training). Neither effect was related to the frequency with which the goal-response was elicited in extinction.

FURTHER VALIDATION OF SUBJECTIVE SCALES FOR LOUDNESS AND BRIGHTNESS BY MEANS OF CROSS-MODALITY MATCHING

By R. T. Root, U. S. Army Personnel Research Office, and
Sherman Ross, American Psychological Association

The past thirty years have seen a renewed interest in the question asked by Fechner regarding the relationship between stimulus and sensation. S. S. Stevens considers this development a "product of a reoriented psychophysics," using "direct" methods of ratio scaling to generate scales of perceptual magnitude.[1] Many perceptual continua have been determined through the use of these methods, among them loudness, brightness, duration, and numerousness.[2] Research results have shown that for prothetic continua (where discrimination presumably involves an additive process and the physical continuum is based upon the intensity of the stimulus, as opposed to metathetic, or substitutive, continua), subjective magnitude grows as a power function of stimulus intensity. Equal stimulus-ratios have been found to result in equal sensation-ratios. Mathematically, this may be expressed as: $S = kR^n$, where n represents the power to which the physical stimulus, R, must be raised to be proportional to the sensation, S.

The prothetic continua of loudness and brightness have been most extensively studied. Research has shown that loudness increases as the 0.30 power of the sound energy (or as the 0.60 power of the sound pressure) of the auditory stimulus, and that brightness increases as the 0.33 power of the luminance (photometric brightness) of the visual stimulus.[3]

Recent research has used the method of cross-modality matching to study the relations between the exponents of the various prothetic continua. Cross-modality matching involves direct matching between the apparent strengths of sensations on pairs of such continua. In the derivation of the subjective magnitude functions

* Received for publication September 5, 1962.

[1] S. S. Stevens, On the psychophysical law, *Psychol. Rev.*, 64, 1957, 153-181.

[2] For a comprehensive review of this research, see S. S. Stevens and E. H. Galanter, Ratio-scales and category-scales for a dozen perceptual continua, *J. exp. Psychol.*, 54, 1957, 377-411.

[3] Stevens, A scale for the measurement of a psychological magnitude: loudness, *Psychol. Rev.*, 43, 1936, 405-416; The direct estimation of sensory magnitudes: Loudness, this Journal, 69, 1956, 1-25; R. M. Hanes, A subjective brightness scale, *J. exp. Psychol.*, 30, 1949, 438-452; Hanes, The construction of subjective brightness scales from fractionation data: a validation, *ibid.*, 719-728.

for perceptual continua, it has been shown that O's could make numerical estimations of their sensations, which when plotted against the physical stimulus continuum, generated the subjective magnitude function. In an analogous fashion, if the O were to estimate his subjective impression of stimuli on one physical continuum by matching them with apparent intensities on another continuum, it would be possible, by this method of magnitude-estimation, to validate the exponents of the two subjective magnitude functions under study.

If values for S (sensation) on two prothetic continua are equated at various intensity levels, the ratio of the two exponents will define the slope of the function generated by the cross-modality comparisons on a log-log plot. The two R values at the various intensity levels should bear the following relation:

$$\log R_1 = n \log R_2/m$$

Cross-modality matching of various continua has shown that the form of the equal-sensation function can be predicted from a knowledge of the two subjective magnitude-functions under study and the slope of this function can, in fact, be predicted from the exponents for the two subjective magnitude-functions.[4]

These studies have shown that the slope of the equal-sensation function is dependent in part upon which of the two stimuli is fixed and which is adjusted. Stevens has called this "regression" or a "centering tendency" and makes the analogue to the two regression-lines in a correlation-plot. The size of this regression-effect appears to depend mostly on the difficulty of the judgment involved. A balanced design can be used, where each stimulus serves as both the standard stimulus and the variable.

The present study was designed to examine the equal-sensation function for loudness and brightness to ascertain whether the slope of this function could be predicted on a knowledge of the two exponents. As the exponent for the loudness-function has been given as 0.30 and for brightness as 0.33, the slope of the equal-sensation function relating intensities on the two continua should be given by the ratio of the two exponents; *i.e.* the slope should approximate 0.91.

Apparatus and procedure. The Os for this experiment were 10 men undergraduates at the University of Maryland. The task for the O involved matching the apparent intensity of auditory stimuli presented by E with visual stimuli under O's control. The auditory stimuli consisted of white noise at the following intensities: 54, 62, 70, 74, 80, 84, and 90 db. A white noise generator was used to produce the auditory stimuli, which were presented through a speaker 2.5 ft. from the Os. The intensity-level was measured with a noise meter placed at the O's

[4] J. C. Stevens and J. D. Mack, Scales of apparent force, *J. exp. Psychol.*, 58, 1959, 405-413: Stevens, On the validity of the loudness scale, *J. acoust. Soc. Amer.*, 31, 1959, 995-1003; Cross-modality validation of subjective scales for loudness, vibration, and electric shock, *J. exp. Psychol.*, 57, 1959, 201-209; S. S. Stevens and J. C. Stevens, Warmth and cold: dynamics of sensory intensity, *op. cit.*, 60, 1960, 183-192.

position. For the presentation of the visual stimuli, experimental apparatus was constructed, consisting of an ordinary 100-w. light bulb placed in a housing 7 × 5 × 4 in., painted with glossy white paint. A red Wratten Filter, No. 25, was placed directly in front of the light-source that the color of the visual stimuli would not change as the voltage decreased. Connected in series with the light-source was a powerstation, which varied voltage from 0 to 135 v. The intensity of the light-source, as the voltage varied from 0 to 135 v., was calibrated in terms of apparent foot-candles with a Macbeth Illuminometer. A rubber hood protruded from the body of the apparatus used to display the visual stimuli. The *O* placed his head against this hood, so for all *O*s the distance from the light-source to the eyes was 12 in.

The *O* was placed in front of the apparatus for presenting the visual stimulus, with his eyes up to the rubber hood. His right hand was placed on the control knob of the powerstation. When presented with an auditory stimulus, *O* turned the control knob until he felt that the brightness of the light-source was of the same

TABLE I

MEDIAN LUMINANCE-SETTING AT EACH NOISE LEVEL

	Noise level (db·re 0.0002 dyne per cm.²)						
	54	62	70	74	80	84	90
Median luminance setting (db·re 10^{-10} lambert)	66.29	69.35	80.83	85.23	87.70	93.15	97.16

magnitude as the loudness of the auditory stimulus. Each *O* was given 28 stimulus-presentations, four at each of the seven sound intensity-levels. The 28 stimulus-presentations were made in random order. For the sample of 10 *O*s there was a total of 40 brightness-settings at each sound-level.

Results. All brightness-settings were converted to *db* re: 10^{-10} lamberts.[5] For each sound-level, the median brightness-setting (in terms of db) was determined. These values are given in Table I. The median settings were then plotted against loudness-levels and the best fitting line through the obtained points was determined by the method of least squares. The results are shown in Fig. 1.

The slope of the line shown in Fig. 1 was found to be 0.908. The slope of the line to be expected from the ratio of the exponents for the subjective magnitude-functions for loudness and brightness was 0.909. It may be seen that the obtained results are extremely close to the expected value.

[5] Conversion of median brightness-settings to db re: 10^{-10} lamberts was made in accordance to the procedure outlined in S. S. Stevens, Decibels of light and sound, *Physics Today*, 8 (10), 1955, 12-17.

Discussion. The results of this experiment show that through the range of intensities used, the brightness-settings of the 10 *O*s were made proportional to the loudness-setting. On a log-log plot the median brightness settings were a linear function of the loudness-settings. The results are consistent with exponents of 0.30 for loudness and 0.33 for brightness. The extremely close correspondence between the expected and observed values for the slope of the equal-sensation function attests to their validity.

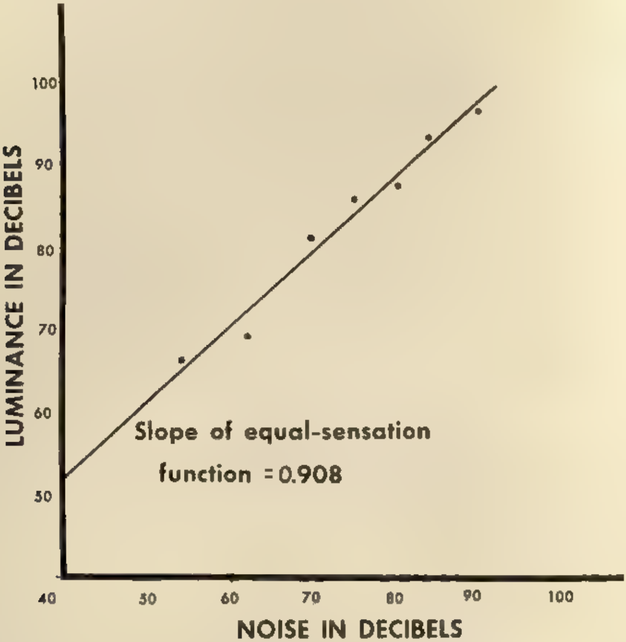

FIG. 1. EQUAL-INTENSITY FUNCTION

It should be noted that the visual stimulus of the present study was *red* light, not white light as used in previous studies. The value of 0.33 for the exponent of the subjective magnitude-function for brightness was obtained, using white light. The only experiment that has studied the subjective brightness of red light was that of Hanes.[6] He found that for red light the over-all slope of the function was slightly less than that for white light. A number of his stimuli were, however, near the threshold for photopic vision. As a result of the Purkinje shift, at low intensity-levels an increase in physical brightness was required to achieve a given level of apparent brightness for the stimuli from the long-wave end of the visible spectrum. The slope of the subjective magnitude-function was thereby decreased. The expected median values for the visual stimuli for this experiment,

[6] R. M. Hanes, A subjective brightness scale, *J. exp. Psychol.*, 30, 1949, 438-452.

however, were not expected to fall near the photopic threshold. Hanes found that for brightness-levels not near the photopic threshold there was a very close approximation to the subjective magnitude-curve for white light. The present study offers further evidence that the subjective magnitude-function for red light is essentially the same as that for white light. In this same context, J. C. Stevens found that the subjective magnitude-functions for white noise and a 1000 ~ tone were essentially identical.[7]

The present design did not permit study of the "regression effect." The *Os* were not asked to match apparent brightness with apparent loudness. In view of the close agreement between the expected and observed slopes and the apparent lack of difficulty in making the judgments involved, however, the slope of the line relating the matching of brightness with loudness would probably not differ to any great degree from that obtained here. We believe, therefore, that the results presented may be thought of as descriptive of the dynamic characteristics of these two sense-modalities.

---

[7] J. C. Stevens and Endel Tulving, Estimations of loudness by a group of untrained observers, this JOURNAL, 70, 1957, 600-605.

DIFFERENCES IN PERCEIVED COLOR AS A FUNCTION OF CHARACTERISTIC COLOR

By JOHN L. DELK and SAMUEL FILLENBAUM,
University of North Carolina

The evidence on the problem of memory-color is far from clear. Duncker[1] and Bruner, Postman, and Rodrigues[2] have shown that, under impoverished or marginal conditions of observation, the apparent color of an object may shift in the direction of its characteristic color. Bolles, Hulicka, and Hanly have suggested, however, that such effects may be obtained only under conditions in which no exact color-match can be made and where, therefore, a true "psychophysical equation is impossible."[3] Since the critical figure and the adjustable color-mixer were not contiguous but viewed successively in these studies, the effects obtained could have been on remembered color rather than on perceived color.

Harper placed the critical figures, one at a time, directly in front of the color-mixer, which served as ground.[4] Using cut-outs of red-associated and neutral objects, he found significantly more red needed to match the former than the latter. Fisher, Hull, and Holtz repeated and extended Harper's study, but were only able in part to replicate his results. Whatever differences were found between red-associated and neutral objects tended to be rather small.[5]

The procedure of the present experiment was patterned after those of the latter two, but there were several changes with regard to experimental arrangements and instructions. In the earlier experiments, the stimulus-figures were very small, but those of the present experiment were considerably larger (by a factor of 5-10). Harper was unable to illuminate his figures evenly, and because of the technique used to mount them, the figures used by Fisher may not have been identical in color. In the present

* Received for publication July 6, 1964.

[1] Karl Duncker, The influence of past experience upon perceptual properties, this JOURNAL, 52, 1939, 255-265.

[2] J. S. Bruner, Leo Postman, and John Rodrigues, Expectation and the perception of color, this JOURNAL, 64, 1951, 216-227.

[3] R. C. Bolles, I. M. Hulicka, and Barbara Hanly, Color judgments as a function of stimulus conditions and memory color, *Canad. J. Psychol.*, 13, 1959, 175-185.

[4] R. S. Harper, The perceptual modification of colored figures, this JOURNAL, 66, 1953, 86-89.

[5] S. C. Fisher, Chester Hull, and Paul Holtz. Past experience and perception: Memory color, this JOURNAL, 69, 1956, 546-560.

experiment, by contrast, there was even illumination for the figures which were all of exactly the same color. Furthermore, the instructions given to the *Ss* of the present experiment were varied systematically to assess the possible biasing effects of the instructions used in the earlier experiments. Three different instructions were employed: (a) instructions similar to those used previously with red-associated figures called "reddish" and the other figures called "yellowish-orange," (b) instructions verbally identifying or naming each figure, as in (a), but with nothing whatsoever said about color, and (c) instructions without any verbal identification or naming of any of the figures. Finally, while Harper and Fisher used only two classes of figures, red-associated figures and abstract neutral figures, a third class—characteristically non-red figures—was used here.

Apparatus and materials. Ten figures were cut from the same sheet of orange-red cardboard—see Munsell chip R/5/12 for approximate color[6]—in the shapes of a heart, an apple, a pair of lips, an oval, a circle, an ellipse, a horse's head (front view), a bell, a mushroom, and a square. These figures, identical in color, were of approximately the same area, ranging from about 28-32 sq. cm. The first three of these figures (heart, apple, and lips) usually are associated with the color red. The second three (oval, circle, and ellipse) usually are associated with no particular color. The third three (horse, bell, and mushroom) usually are associated with colors other than red. The tenth figure (square) was used only as a practice or demonstrational figure and was not included in the experiment proper.

A differential color-mixer which permits the mixing of two colors to produce continuously varying intermediate shades was used. The mixture could be varied from 360° red (Munsell chip R/3/8) and 0° yellow-orange (Munsell chip YR/6/10) to 0° red and 360° yellow-orange. Each of the figures could be mounted directly in front of the center of the color-mixer, which thus served as ground for each figure.

The color-mixer was placed upon a table in the corner of the experimental room. About 6.5 in. in front of the color-mixer was a black screen, 2 × 4 ft., in which there was a 5.5-in. aperture. On the inside of the screen, over the aperture, a sheet of waxed paper was pasted to reduce visual acuity and to obtain a better blend of figure and ground. Both the figure and ground were illuminated by a 9-in. circular fluorescent lamp so attached to the inside of the screen that it formed a circle around the aperture. This lighting arrangement provided even, soft illumination of both figure and ground, and cast no shadows of the figure on the ground. The overhead lights in the room were turned off and the window shade drawn, leaving the room quite dimly illuminated. From his seated position, *S* could see only the black screen, the aperture, and, through the aperture, the figure upon a field of orange-red.

Subjects. The *Ss* of this experiment were 60 undergraduate men at the University of North Carolina, recruited from a course in introductory psychology.

[6] In the Munsell notation, the letter indicates hue, the first number indicates value (saturation), and the second number indicates chroma (intensity or strength). See A. H. Munsell, *The Atlas of the Munsell Color System,* 1929. Two *Os* independently matched the orange-red cardboard to the same Munsell chip.

Procedure. The *Ss* were randomly assigned to the three instructional conditions, with 20 *Ss* serving in each condition. All were told that they would see a number of figures, one at a time, against a background that could be changed from a light red to a dark red, or *vice versa.* They were to instruct *E* to make the background darker or lighter until it was the same color as the figure, *i.e.* until the figure could no longer be distinguished from the background. For *Ss* in Group I (the Color-Label Group), as each figure appeared it was called a "yellowish-orange horse," a "yellowish-orange oval," and so forth, except for the three red-associated forms which were described as a "reddish heart," a "reddish apple," and a "reddish pair of lips." The *Ss* in Group II (the Label-Only Group) were told the name of each figure as it was presented, but there was no reference to color. The

TABLE I

MEAN SETTING (IN DEGREES OF RED) FOR EACH FIGURE UNDER EACH INSTRUCTIONAL CONDITION

	Figure											
Group	red-associated				neutral				non-red			
	heart	apple	lips	combined	oval	circle	ellipse	combined	horse	bell	mushroom	combined
Color-Label	265	284	272	273	241	244	243	243	243	241	228	237
Label-Only	263	271	255	263	240	232	238	237	237	226	225	230
Self-Labeling	269	280	254	268	245	236	248	243	248	231	230	237
means				268				241				235

Ss of Group III (the Self-Labelling Group) were not instructed either as to the name or the color of any figure, but were asked to tell *E* the name of each figure.

Two matches (one ascending trial and one descending trial) were obtained for each figure. For every *S* in each condition, the square was presented first as a practice-figure. The remaining nine figures then were presented in a randomized order. The random orders of presentation used with Group I were used also for the other two groups.

Results. Table I presents the mean settings for each figure and each group. Analysis of variance reveals only one significant effect, that of Figures, with $F = 56.18$, $df = 8/456$, $p < 0.01$. Although differences in instruction had no appreciable effect on the judgments, there were considerable differences between the figures, with red-associated figures requiring more red in the ground for the match than neutral or non-red-associated figures; there was no interaction between Instructions and Figures.

Table II contrasts the various means.[7] It reveals that each of the three red-associated figures was seen as significantly more red than any of the other six figures, and that the apple was seen as significantly more red than the heart or lips. The mean differences between the red-associated

[7] The analysis was based on procedures outlined in B. J. Winer, *Statistical Principles in Experimental Design*, 1962, 80-85.

figures and the other figures are not only highly significant statistically, but also of substantial magnitude perceptually. This becomes evident if one compares the Munsell chips which approximately match (to the eyes of two independent judges) the mean settings for the three classes of figure. For red-associated figures, the Munsell chip is R/4/12; for the neutral figures, R/5/12; and for the non-red figures, R/5/13. The difference between chip R/4/12 and the other two chips is considerable. The other two chips are also perceptually distinct, although the difference between them is smaller.

TABLE II

TESTS OF MEAN-DIFFERENCES BY THE NEWMAN-KEULS PROCEDURE

Figure: Ordered Means:		Mushroom 1 227.8	Bell 2 232.1	Circle 3 237.0	Oval 4 242.0	Ellipse 5 242.6	Horse 6 242.7	Lips 7 259.9	Heart 8 265.4	Apple 9 278.7
Differences between pairs	Mushroom		4.3	9.2*	14.2†	14.8†	14.9†	32.1†	37.6†	50.9†
	Bell			4.9	9.9*	10.5*	10.6*	27.8†	33.3†	46.6†
	Circle				5.0	5.6	5.7	22.9†	28.4†	41.7†
	Oval					0.6	0.7	17.9†	23.4†	36.7†
	Ellipse						0.1	17.3†	22.8†	36.1†
	Horse							17.2†	22.7†	36.0†
	Lips								5.5	18.8†
	Heart									13.3†
q critical at 5% level:			6.27	7.53	8.27	8.78	9.18	9.50	9.77	10.04
q critical at 1% level:			8.29	9.41	10.08	10.55	10.91	11.22	11.47	11.67

* Significant at the 5% level.
† Significant at the 1% level.

Why were the identically colored figures responded to as though they were actually different in color? All figures were cut from the same colored paper; were of approximately the same area (the correlation between setting and area was only -0.05); had relatively smooth contours; and were viewed under the same conditions of illumination, against the same ground. One is led to the conclusion that past association of color and form does in some way influence perceived color, since that is the one respect in which the figures did clearly differ. To interpret the results in terms of memory-color does not, of course, explain them, since little can be said at this time as to the mechanism by which an interaction of sensory and associative factors might occur.

Summary. This experiment had two purposes: (1) to investigate the possible effect of characteristic color upon perceived color; and (2) to assess the possible biasing effects of instructions on color-matching judgments. Although differences in instruction had no effect upon the judgments, there was a highly significant effect attributable to characteristic color. Each of the three red-associated figures was seen as redder than each of the other six figures (three neutral figures and three non-red figures). There was no interaction of characteristic color with instructions.

VERBALIZATION AS A STIMULUS-TRACE IN *GSR* CONDITIONING

By Harold W. Coppock, Arizona State University

Trace-conditioning has been too difficult to establish and too vulnerable to "disinhibition" to provide a plausible explanation for acquisition of phobic reactions to stimuli antedating the trauma by more than a few seconds. Few uninstitutionalized individuals have as well controlled and as constant an environment as the one provided for Pavlov's dogs. Stimulus-traces have been inferred not only from trace and other forms of conditioning and reaction-time data, but also from such phenomena as time error, set, sensitization and delayed responding.[1] Trace or set disappears following external disinhibiting stimuli, running more than 18 in. in a runway, turning a wheel in a box, and following almost any major postural readjustment of animals simpler than the primate.[2]

The present laboratory experiments in conditioning at long *CS-US* intervals attempt to increase the constancy of environmental stimuli and also the magnitude of the stimulus-trace by instructing *S* to verbalize. The instructions are designed to control stimulus-traces which probably comprise a large fraction of the stimulus-population of man: the sound of his own vocal talking.

In *GSR* conditioning, rate of increase in habit-strength is considered to be an increasing function of small magnitudes of *CS*-trace remaining at the time of the *US*, relative to the magnitude of the irrelevant stimuli.[3] This interpretation is made in spite of recently published lack of statisti-

---

* Received for publication April 10, 1961.

[1] Bradley Reynolds, The acquisition of a trace conditioned response as a function of the magnitude of the stimulus trace, *J. exp. Psychol.*, 35, 1945, 15-30, 81-95; R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, 1954.

[2] M. O. Wilson, Symbolic behavior in the white rat: II. Relation to quality of interpolated activity to adequacy of the delayed responses, *J. comp. physiol. Psychol.*, 18, 1934, 367-384; D. R. Meyer, Chungsoo Cho, A. F. Wesemann, On problems of conditioning discriminated lever-press avoidance responses, *Psychol. Rev.*, 67, 1960, 224-228.

[3] The ineffectiveness of an *intense CS* (H. D. Kimmel, Amount of conditioning and intensity of conditioned stimulus, *J. exp. Psychol.*, 58, 1959, 283-288.) *i.e.* a *CS* which might elicit larger amplitude of *GSR* than the so-called *US*, is attributed to backward conditioning which has been obscured by inappropriate classification of stimuli. See also, G. H. S. Razran, The dominance-contiguity theory of the acquisition of classical conditioning, *Psychol. Bull.*, 54, 1957, 1-46.

cal significance between slight variations in clearly supra-threshold *CS* intensities in constant laboratory environments of men and of rats.[4]

EXPERIMENT I

Method. Each *S* was placed on a contour chair in a darkened room. Since extraneous light could be controlled more readily than extraneous sound, a visual *CS* was employed: onset of illumination of the wall in front of *S*. This occurred at 60-sec. intervals and was terminated after 30 sec. The unconditioned stimulus was a doorbell that started 2 sec. after *CS* and lasted for 2 sec. The bell was pressed against *S*'s head. An ohmmeter, previously described,[5] continuously recorded palmar skin-resistance with pen excursions proportional to percentage-change in resistance. *GSR*s to the *CS* were determined by taking the first decrease in resistance starting from 1 to 4 sec. after the *CS*. These limits seemed to include all responses around the time of the *CS*. Most *GSR*s started after the usual latency of 2 sec.

Data from 48 *Ss* were used to compare the amount of 2-sec. forward conditioning obtained under two instructions: (a) to talk freely about the experiment; and (b) to make successive audible subtractions of 3. The *CS* was presented by *E* at 1-min. intervals for five adaptation-trials, paired with the *US* for two conditioning trials, and presented alone for three extinction trials. An increase in amplitude of responses from the last adaptation-trial to the first extinction-trial was the measure of conditioning.

Results. Group means indicated no conditioning occurred with either instruction. The number of conditioning trials was fewer than customary for conditioning motor responses, yet comparable to the number employed in successful *GSR* conditioning with weak *CS* and *US*.[6] Unsystematic observations suggested that *Ss* were so naïve about conditioning and so impressed with the setting, that talking about the experiment introduced as much variability as the subtraction.

EXPERIMENT II

Method. Experiment II was designed to increase consistency in the verbalization and to try shorter *CS-US* intervals. As in Experiment I, five adaptation-trials, two conditioning, and three extinction-trials were given. In the first treatment, 12 *Ss* were instructed to count after each *CS* as follows, "one thousand-one, one thousand-two." The *US* was presented 0.5 sec. after *S* finished saying "one thousand two." These instructions were repeated during preliminary timed practice to insure uni-

[4] D. A. Grant and D. E. Schneider, Intensity of the conditioned stimulus and strength of conditioning: II. The conditioned *GSR* to an auditory stimulus, *J. exp. Psychol.*, 39, 1949, 35-40; W. J. Hansche and D. A. Grant, Onset versus termination of a stimulus as the *CS* in eyelid conditioning, *J. exp. Psychol.*, 59, 1960, 19-26; W. C. Miller and J. E. Greene, Generalization of an avoidance response to varying intensities of sound, *J. comp. physiol. Psychol.*, 47, 1954, 136-139.

[5] H. W. Coppock, Responses of subjects to their own galvanic skin responses, *J. abnorm. soc. Psychol.*, 50, 1955, 25-28.

[6] H. W. Coppock and R. M. Chambers, GSR conditioning: An illustration of useless distinctions between types of conditioning, *Psychol. Rep.*, 5, 1959, 171-177.

formity of rates of counting. In each of the second and third treatments 12 *Ss* were instructed to sit quietly and relax. The *CS-US* interval was 0.5 sec. in the second treatment and 2 sec. in the third. In each treatment-group, the *CS* was light-onset for half of the *Ss* and light-termination for the other half; light being on for 30 sec. and off for 30 sec., as in Experiment I.

Results. Instructions to count yielded no significant change in *GSR*, using a *t*-test for related measures. As a matter of fact, the responses of 5 of the 12 *Ss* in the first group became smaller rather than larger. The relaxing instructions produced significant conditioning not only with the 0.5-sec. interval ($p < 0.015$), but also with the 2-sec. interval ($p < 0.015$). Contrary to expectations, analysis of variance of the 2-sec. groups (the only groups with homogeneous variance) indicated that the instructions to relax produced significantly more conditioning than the instructions to count ($p < 0.02$). Counting apparently interfered equally with both types of *CS*, since interaction is *not* significant. The same analysis provided the additional information that conditioning, with *CS* being onset was more effective than conditioning with *CS* being offset ($p < 0.04$). Although this is not consistent with recent data on eyelid-conditioning, it is consistent with our expectation, in view of the greater magnitude of sensory output after onset than after termination of visual stimulation.[7]

Significant conditioning, when the *CS* preceded shock by 2 sec. is not, at first glance, consistent with the report of White and Schlosberg that "nothing but pseudo-conditioning is obtained with the *GSR* if the interval of delay is 1 sec. or more."[8] Our *CS* was produced, however, by a 250-w. photoflood lamp which illuminated the wall in front of *S*. It is reasonable to suppose that this elicited a larger and hence longer-lasting trace than Schlosberg's *CS* from a 12-w. bulb behind a piece of milk glass. The suggestion that the maximal effective period of delay between *CS* and *US* is an increasing function of the intensity of *CS* is, however, consistent with the observations, in traumatic avoidance conditioning, of conditioned autonomic reactions even when the *CS-US* interval was 10 sec.[9] Their *CS* consisted of the termination of 80-w. illumination directly over the dog plus the raising, probably not without noise, of a partition.

In our experiment, the reduction in effectiveness of conditioning due to *S*'s chain of verbal responses is not attributed to above-optimal auditory *CS* intensity, since *S*'s voice was probably not above 35 db., which was an effective *CS* intensity for

[7] Hansche and Grant, *op. cit.*, 26; S. S. Stevens, *Handbook of Experimented Psychology*, 1951, 966.

[8] C. T. White and Harold Schlosberg, Degree of conditioning of the GSR as a function of the period of delay, *J. exp. Psychol.*, 43, 1952, 357-362.

[9] R. L. Solomon, L. J. Kamin, and L. C. Wynne, Traumatic avoidance learning: the outcomes of several extinction procedures with dogs, *J. abnorm. soc. Psychol.*, 48, 1953, 291-302.

Kimmel.[10] It might indicate that any verbalization by *S* interferes with the visual trace-process. For *CS-US* delays as short as 2 sec., there may be some more effective, autonomous physiological trace such as an after-image or an eye-movement sequence which consistently follows the visual *CS*. This trace would mediate conditioning under instructions to relax, but might be disrupted by instructions to count. Perhaps in conditioning as in retroactive inhibition, "summation of stimuli" depends on compatibility of responses.[11]

Summary

(1) Galvanic skin-responses, conditioned in three trials, with weak *CS* and *US*, is greater if the onset of illumination rather than its termination is the effective *CS*. This is attributed to the greater magnitude of the sensory trace following onset.

(2) Conditioning with a 0.5-sec. interval between *CS* and *US* did not occur after instructing *S* to verbalize in a manner which is unsystematic with respect to the occurrences of the *US*.

(3) Conditioning with a 2.0-sec. interval was reduced by instructing *S* to start counting seconds at each *CS*, and by presenting *US* after the response "two" on each conditioning trial. This result contradicts our deductions from the theory of stimulus-sampling and from the principle of summation of stimuli.

---

[10] H. D. Kimmel, *op. cit.*, 183.
[11] Woodworth and Schlosberg, *op. cit.*, 752.

ON THE MECHANISM OF TONAL CHROMA IN ABSOLUTE PITCH

By Herbert S. Korpell, Cornell University

Bachem and Révész have suggested that two factors are involved in 'absolute pitch,' which is the ability to recognize and define, by naming or singing, a musical pitch without the aid of a reference-tone. These factors are recognition of 'tone height,' roughly equivalent to frequency, and of 'tone chroma,' which renders each note of the scale phenomenally distinct.[1] It may be that the information essential to the perception of chroma resides in the sound delivered to the ear, or this information may arise within the ear itself. If it is contained in the distal stimulus, it must result from the distribution of accompanying overtones, which are known to vary from note to note on a given instrument.[2] If the stimulus proximal to the auditory nerves creates the subjective phenomenon of chroma, it must be as a result of the non-linear transmission of the ear-mechanism.[3]

Musical tones were re-recorded slightly higher or lower than their original pitch, without changing the proportions of their component harmonics. If an unique pattern of overtones for each note is responsible for the perception of chroma, the *Os* would be expected to identify these altered notes with the originally recorded letter-names; if chroma arises within the mechanism of the ear, the *Os* should name the transposed notes at their new frequencies.

Observers. The *Os* were 23 volunteers demonstrating pitch-naming ability. All were associated with either Cornell University or Ithaca College.[4]

---

* Received for publication July 7, 1964. The author is indebted to Professors Julian Hochberg, Clyde Ingalls, Robert Stine, and Howard Simmons, and to Mr. Horace Reynolds, for advice and technical assistance.

[1] Albert Bachem, Tone height and tone chroma as two different pitch qualities, *Acta Psychol.*, 7, 1950, 80-88; Géza Révész, *Introduction to the Psychology of Music*, 1954, 56-69.

[2] See, for example, James Jeans, *Science and Music*, 1953, 95; John Backus, Vibrations of the reed and the air column in the clarinet, *J. acoust. Soc. Am.*, 33, 1961, 806-9.

[3] It is possible that a higher-order neural variable not presently understood is involved, but a reasonable basic explanation largely avoiding such recourse already has been presented (Bachem, Various types of absolute pitch, *J. acoust. Soc. Am.*, 9, 1937, 150-1).

[4] The *Os* did not prove of equal ability, but uncontroversial criteria for defining absolute pitch have not been presented in the literature.

Apparatus. Notes played by competent musicians were recorded with a sound-system consisting of a Shure 55S dynamic microphone, an Ampex 354 tape-deck, and an Altec 1568A amplifier, and with a Wollensak T1500 tape-recorder. Transposition was accomplished by taping directly from the Wollensak into an International Radio and Electronics 'Crown' recorder, the speed of the Wollensak governed by feeding the powering alternating current through a Hewlett Packard oscillator 207A and two 'McIntosh 60' amplifiers. The *Os* were tested with the Wollensak. Scotch '120' and Kodak '31A' 1½-mm. triacetate tape was run at 7½ in./sec. A Kay Electric "Missilizer Missile Data-Reduction Spectrograph" was used for overtone-analysis. Stimulus-tones were accurate to within 0.10 semitones,

TABLE I

APPROXIMATE DURATION OF ELEMENTS OF THE COMPLETE TEST-TAPE
(In minutes and seconds)

The 30 tones in each part were separated by pauses, and, in Part III, by a disturbance following the pause. Disturbances, consisting of randomly ordered presentations of white noise and singing and instrumental complexes similar to the 'tuning up' of an orchestra, were included, but they are not significant to the present discussion. Pause between parts, 0:10; total length of tape, 14:45.

	Part I	Part II	Part III	Part IV
Tones	0:01	0:02	0:02	0:02
Pauses	0:02	0:04	0:05	0:04
Disturbances	—	—	0:07	—
Totals	1:28	2:56	6:55	2:56

as tested by matching them with the signal from a Hewlett Packard function-generator 202A on a Dumont 304H oscillograph.

Procedure. The program consisted of 120 musical tones divided equally into four parts, as listed in Table I. Pitches, given in random order, were taken from the middle two octaves of the ranges of flute, violin, men's and women's voices, saxophone, baritone horn, and piano. Each altered note was raised or lowered by one or two semitones. Disregarding pure tones produced by a tone-generator and notes left unaltered in the complete experimental program, 56 stimuli are pertinent to the present discussion. The *Os*, tested individually, were asked to write the letter-names of the notes as they heard them, according to the method of single-stimulus absolute judgment.

Results and conclusions. The total score for each *O* was calculated in two ways: counting a response correct if it was the same as the note originally recorded, and counting a response correct if it was the same as the note at its new frequency. In the results of the present experiment, notes were identified according to the new frequencies, beyond the 0.1% level of confidence. It is concluded, therefore, that the criteria for chroma arise within the mechanism of the ear.

An analysis of overtones supports this conclusion. Fig. 1 shows gross differences in the proportions of harmonics of altered and unaltered stimulus-tones played by one instrument which were presented at the same frequency (*G* one-and-a-half octaves above piano's middle *C*). There

is also a large difference between two tones originally played at the same pitch (F), presumably due to minute differences in the strength and direction of the air-stream impinging on the lip plate of the flute. Since the tones represented by these greatly different spectrographs yield sensations of identical chroma, it is unlikely that such chroma is contained in the specific overtone-structure.

This conclusion finds application in research on absolute pitch. It suggests, for example, that when O becomes aware that a phonograph turntable is not operating at proper speed, he does so because of cues such as familiarity with the proper key and tempo of the piece being played, timbre of familiar instruments on certain notes, noise from the phonograph equipment, and so forth, rather than because of absolute-pitch chroma.

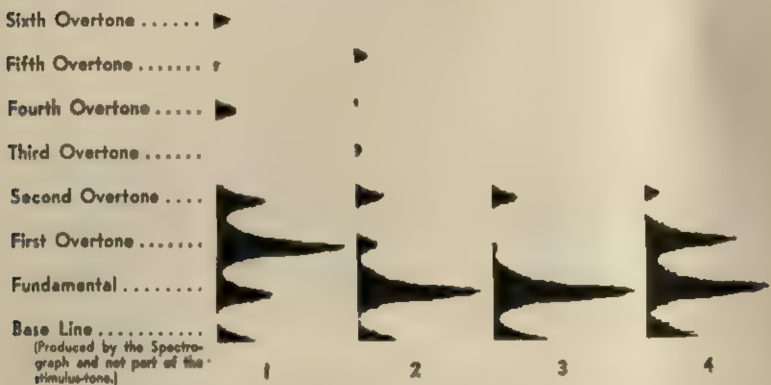

FIG. 1. SPECTROGRAPHIC ANALYSIS OF FOUR FLUTE-TONES OF IDENTICAL PITCH AND CHROMA

(1) An unaltered G; (2) A G originally recorded as an A two semitones above its present frequency; (3) A G which was originally an F two semitones down; (4) Another G which was originally an F.)

Summary. Observers adept at the identification of absolute pitch were presented musical tones which had been re-recorded at a speed different from the original recording. These notes of new frequencies retained their original overtone characteristics. Since tone chroma, the quality by which each note of the scale may be identified as phenomenally distinct, must arise in either the proximal or the distal stimulus, and since the Os identified the tones correctly according to the frequency at which they were played as test-stimuli, and not according to the overtone structure which corresponded to a note of a different frequency, it is concluded that chroma arises as a result of the non-linear response of the ear.

PERCEPTUAL INVARIANCE IN THE KINETIC DEPTH-EFFECT

WILLIAM EPSTEIN, University of Kansas

Under appropriate conditions, a continuous sequence of shadow-transformations will yield the appearance of a rigid form, having internal depth and turning in space. This is the kinetic depth-effect (*KDE*) studied by Wallach and O'Connell.[1] Wallach's experiments and the related literature have been reviewed by Braunstein.[2] A peripheral result reported by Wallach and O'Connell provided the starting-point for the present studies.[3] They found that *O*s could judge the internal depth of the shadow-casting figure with considerable accuracy. The main objective of the present research was to learn whether *both* internal depth and amount of turning can be accurately judged under varying lengths of transformation in the *KDE*-situation. Accordingly, judgments of depth and turning were obtained from the same *O*s.

Apparatus. A device similar to the shadow-transformers used by Gibson and Gibson,[4] and by Wallach and O'Connell[5] was constructed. The apparatus was designed to permit continuous variations of the arc and rate of oscillation of the shadow-casting form. In addition to the *KDE*-apparatus, a protractor-like device was constructed for the purpose of obtaining measurements of the perceived turning.

Materials. The shadow-casting figure was adapted from Wallach and O'Connell.[6] It was a wire parallelogram containing one diagonal, which was so bent along this diagonal that the planes of the upper and lower triangles formed an angle of 120° with each other. The form was mounted on a vertical rod which was its axis of oscillation.

A series of seven comparison-forms was constructed of wire and painted black. The comparison-forms were bent parallelograms ranging from 90°–150° in steps of 10°. They were mounted on vertical axes and displayed together. The positions of the comparison-forms on the display-board were determined randomly. Three different arrangements were used.

* Received for publication October 18, 1963. This investigation was supported by Grant MH 4153-04 from the U. S. Public Health Service.

[1] Hans Wallach and D. N. O'Connell, The kinetic depth effect, *J. exp. Psychol.*, 45, 1953, 205-217.

[2] M. L. Braunstein, The perception of depth through motion, *Psychol. Bull.*, 59, 1962, 422-433.

[3] Wallach and O'Connell, *op. cit.*, 211-212.

[4] J. J. Gibson and E. J. Gibson, Continuous perspective transformations and the perception of rigid motion, *J. exp. Psychol.*, 54, 1957, 129-138.

[5] Wallach and O'Connell, *op. cit.*, 208.

[6] Wallach and O'Connell, *op. cit.*, 205-217.

Observers. The *O*s were 75 undergraduate students. They were assigned in their order of appearance to the five experimental conditions.

Procedure. The shadow-transformations were produced by oscillating the 120° parallelogram through one of the following five arcs: 15°, 25°, 45°, 65° and 85°. Each *O* made judgments for one arc of oscillation only. Observation was binocular, and no effort was made to restrict *O*, *e.g.* to discourage head-movements. *O* began the experiment by reporting his impressions of a static shadow-form. Then there were four exposures of the shadow-transformations. On two trials, *O* reported his judgments of form by selecting a member of the comparison-series. On the remaining trials, *O* attempted to duplicate the arc of turning of the shadow-casting form by turning his selected form through the desired arc on the protractor-device. The instructions in both cases stressed *objective* matches.

Results. The initial static presentation yielded a three-dimensional impression for 32% of the *O*s. This proportion was increased to 100% by the onset of turning. The judgments of internal depth were highly accurate. The deviations from complete accuracy ranged from 1.34° to 5.00°. There were no significant differences in judgment for the various arcs of oscillation.

The mean judgments of turning were 20°, 31°, 48°, 58°, and 76° for the groups assigned to the 15°, 25°, 45°, 65°, and 85° arcs of oscillation, respectively. Analysis of variance yielded an F of 64.46 ($P < 0.001$, $df = 4/70$). Duncan's range-test showed that all the differences between individual means were significant at the 1% level. An analysis also was made of the accuracy of turning judged in relation to the arc of oscillation; it yielded an F of 2.44, which is not significant.

Discussion. The results show that *both* internal depth and amount of turning can be judged with relative accuracy. This finding conforms to the expectations derived from Gibson's theory,[7] and it is compatible with the shape-slant invariance-hypothesis. With the exception of Langdon's studies,[8] that hypothesis has been tested under conditions of static stimulation, an atypical condition of stimulation which may not reflect the capacity of the perceptual system for partialling out the stimulus-invariants. This may account for the equivocal and contradictory nature of much of the evidence reported by most investigators of the hypothesis.[9] It may be

[7] J. J. Gibson, Perception as a function of stimulation, in Sigmund Koch (ed.), *Psychology: A Study of a Science*, 1, 1958, 456-501.

[8] James Langdon, The perception of changing, *Quart. J. exp. Psychol.*, 3, 1951, 157-165; Further studies of the perception of changing shape, *ibid.*, 5, 1953, 89-107.

[9] These studies are reviewed by William Epstein and J. N. Park (Shape constancy: Functional relationships and theoretical formulations, *Psychol. Bull.*, 60, 1963, 265-288).

that the invariance-hypothesis is more appropriately tested under conditions in which either *O* or the target is in motion.

The finding that the length of the transformation-sequence does not affect the judgment of internal depth was unexpected. Apparently, the shortest sequence (15°) was sufficient to give the depth of the figure. It would be interesting to determine the minimal sequence which would serve this purpose.

SKIN-CONDUCTANCE, ALPHA-ACTIVITY, AND VIGILANCE

By D. R. DAVIES, University of Leicester, and
ANDJELKO KRKOVIC, University of Zagreb

In the typical vigilance-situation, *S* is required to detect and to respond to slight and infrequent changes in stimulation over long periods of time. The usual finding has been that performance declines markedly after about a half-hour at the task.[1] Little attention has thus far been paid to the possibility of a relationship between performance in vigilance-tasks and electrophysiological measures of activation, such as skin-conductance and *EEG* alpha-activity. The present experiment was designed to investigate such a possibility.

Method. The apparatus used to record skin-resistance was a Fels Dermohmeter, Model 22A. The zinc-zinc sulphate paste electrodes were attached to the soles of the feet by means of plastic clamps. Readings to the nearest 1000 Ω were taken at the end of each minute and later translated into conductance-units.

EEG-potentials were recorded from two chlorided silver-disk electrodes secured to the scalp by collodion and placed bilaterally over the occipital lobes about 1 in. above the occipital bulge. Two earlobe-electrodes were used for grounding, and no deflection in amplitude was observed on this channel. Conventional *EEG*-records were obtained by using a Grass Electroencephalograph, Model 111d. From the later stages of amplification of this instrument, the signal was fed into a band-pass electronic filter, rectified in an integrator preamplifier, Model 5U-IS, integrated in a unit-integrator, Model UI-I, and then further amplified and recorded with a polygraph recorder, Model 5. The characteristics of the filter were as follows: maximal relative transmission at 10 c.p.s., with 50% cut-off points, for sine-wave input at the electroencephalograph, at 7.3 and 12.9 c.p.s. The integrator, which provided mathematical integration (area under the curve), was reset each minute by an external timer. The 1-min. values were read to the nearest millimeter from the graphic record, and mean values later were computed for each 15-min. period.

The vigilance-task involved the auditory presentation by tape-recording of a series of digits (1-9) in random order at the rate of 1 per sec. The signal to be detected consisted of a series of three successive odd digits, all of which were different. There were 36 such signals to be detected during the 90-min. session. In each 15-min. period, there were six signals, occurring at 19 sec., 1 min. 28 sec., 3 min. 28 sec., 6 min. 27 sec., 8 min. 25 sec., and 14 min. 58 sec. This cycle, but

* Received for publication April 6, 1964. This research was done at Lehigh University. The authors are grateful to Dr. Josef Brožek for advice and to A. Crowell and R. J. Cunitz for technical assistance.

[1] N. H. Mackworth, Researches in the measurement of human performance, *Med. Res. Council*, Spec. Rep. Ser., No. 268, 1950, 1-156.

not the actual digits, was repeated during the remaining five 15-min. periods. When *S* thought he had detected a signal, he pressed a key which activated a pen of an Esterline-Angus event-recorder.

The *Ss* were 10 college men, 19-23 yr. old, who were paid for their services. They were tested from 1-4 P.M. and performed the task with eyes closed, lying on a bed in an electrically shielded, sound-deadened room. This reclining position was chosen to facilitate relaxation and to minimize the possibility of contaminating the *EEG*-record with action-potentials. A 6-v., 40-w. lamp above the bed provided the illumination. The tape-recorded task was relayed to *S* over an intercommunication system connecting *S*'s and *E*'s rooms. A 15-min. resting record for each electrophysiological measure was obtained both before and after the task was pre-

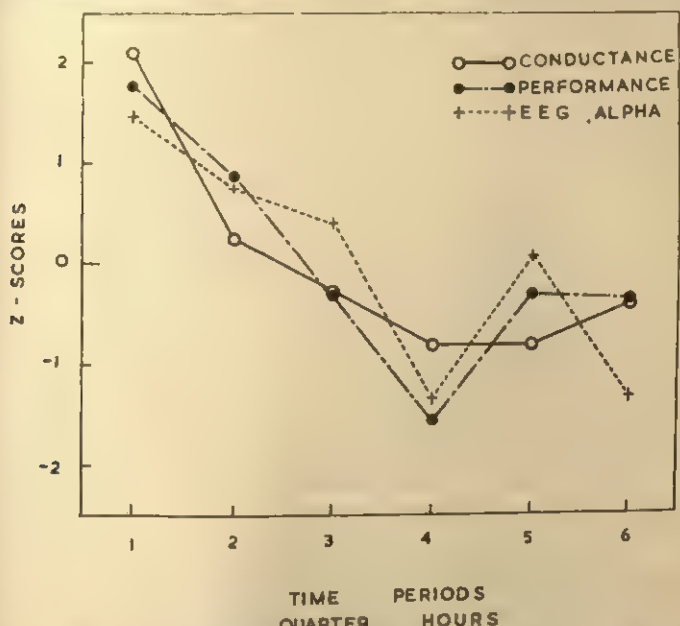

FIG. 1. Z-SCORES FOR CONDUCTANCE PERFORMANCE AND *EEG* ALPHA-ACTIVITY OVER THE SIX QUARTER-HOUR PERIODS OF THE TASK

sented. At the conclusion of the experiment the *Ss* were questioned about their experiences.

Results. Skin-conductance and alpha-activity measures were converted into working-resting ratios for each *S* and the values averaged for all the *Ss* for each of the six 15-min. periods of the task. These ratios (of work to pre-work measures) were computed in order to see more clearly the effect of the task upon the level of activation. For purposes of comparison, the ratios were converted to *Z*-scores, which are plotted in Fig.

1, along with the detection-scores which also were converted to Z-Scores. The three curves are very much the same, all showing a negatively accelerated decline with time. A trend-analysis ($F = 0.86$, $p > 0.50$) revealed no significant difference among the trend-lines of the three measures.

The results of the questionary may be summarized as follows: all the *S*s underestimated the number of signals in the task; the mean estimated number of signals was 27.50, while the actual number was 36. All the *S*s underestimated their own performances, averaging 2.50 signals below their actual scores. Seven of the 10 *S*s reported some daydreaming, and eight reported feelings of drowsiness. When questioned further, these *S*s specified that they felt most drowsy after about an hour at the task. Since they had no means of keeping track of time, it is interesting to note that this estimate agrees well with a finding of slower *EEG*-activity (theta, flattened record) in the fourth 15-min. period of the raw tracing. No expressions of irritability were noted, nor did the *S*s appear to find the task difficult. The most common word used to describe it was "boring."

The results of this experiment lend support to an activational theory of vigilance, which, on the basis of experiments on sensory deprivation, has been used to explain the decline in performance with time.[2] This explanation sometimes has been made in terms of the need for a lively perpetually variable environment to conserve a sufficient level of arousal or activation in the central nervous system.[3] Since it has been demonstrated that alpha-activity and skin-conductance decrease during sleep,[4] it might be expected that decreases in these measures also would occur during vigilance-tasks, in which reports of drowsiness are common.[5]

Summary. An experiment is reported in which the relationship between skin-conductance, alpha-activity, and auditory vigilance was studied. A trend-analysis shows that these three measures appear to be associated over a 90-min. vigil. The results seem to indicate that during the course of a vigil *S* becomes drowsy, as shown by the decrease in *EEG* alpha-activity, skin-conductance, and the level of performance. It is also confirmed by *S*'s replies to a questionary. It is suggested that this drowsiness is produced by the lack of varied stimulation from the task and from the environment.

[2] D. E. Broadbent, *Perception and Communication*, 1958, 1-338.

[3] Donald Wallis and J. A. Samuel, Some experimental studies of radar operating, *Ergonomics*, 4, 1961, 155-168.

[4] R. B. Malmo, Activation, a neuropsychological dimension, *Psychol. Rev.*, 66, 1959, 367-386.

[5] Paul Bakan, Extraversion-introversion and improvement in an auditory vigilance task, *Brit. J. Psychol.*, 50, 1959, 325-332.

EVOCATION AND LEARNING TESTS FOR SECONDARY REINFORCEMENT

By F. J. McGUIGAN, BARBARA EQUEN, and JUDITH STELLINGS,
Hollins College

McGuigan and Crockett trained rats to run to a positive goal (black or white) which contained food.[1] The rats were then tested in a (different) two-choice maze where the goals (black or white) were visible from the choice-point. One result was that the rats ran to the positive goal significantly more often than chance. Clearly the positive stimulus had acquired evocative properties. It is widely assumed, however, that stimuli that have been associated with reinforcement acquire reinforcing properties as well (reinforcing in the sense that the stimulus can increase the probability of a new response) and further that these two acquired functions of stimuli are perfectly correlated. The present study was designed to confirm this conclusion—that stimuli associated with reinforcement (as specified in the McGuigan and Crockett experiment) will not only evoke a response, but will also increase the frequency of a new response, both in a different situation. While positive effects have been obtained in a number of experiments that used either evocation or new learning as a test, apparently both criteria have never been used in the same experiment, a condition necessary for the clearest demonstration.

Method. The *Ss*, 32 Sprague-Dawley female albino rats about 100 days old at the start of the experiment, were reinforced in the presence of white or black in a Grice-type maze. This training was essentially the same as previously reported, except that *Ss* were given 84 training trials.[2] A T-maze, with guillotine type doors immediately after the choice-points and at the entrance to the goal boxes (white and black), was used for the test-trials. All *Ss* were given 16 rotated test-trials and were detained for approximately 5 sec. in the box to which they ran. The empty food-cup from which *Ss* ate during training trials was always in the positive goal-box. Half of the black-positive-trained *Ss* and half of the white-positive *Ss* were assigned, at random, to each of two groups. Group I was tested for evocation, the side on which the positive box appeared being randomly determined on each trial with the restriction that it appeared on each side an equal number of times. Baseboards were placed at the choice-point and were of the same color as the goal-

* Received for publication December 26, 1961.

[1] F. J. McGuigan, and Frances Crockett, Evidence that the secondary reinforcing stimulus must be discriminated, *J. exp. Psychol.*, 55, 1958, 184-187.

[2] McGuigan and Crockett, *op. cit.*, 184-185.

box to which they led. Group II was tested for new learning and hence did not have the baseboards present. Instead, curtains immediately before the goal-box doors hid the color of the goal-boxes until one of them was entered. The technique to control position-habits reported by D'Amato was used for this group; with the exception of the first trial, then, the positive goal-box was always on the same side for a given S.[3] In short, the positive stimulus was in view for the test of evocation, this measure being primarily one of determining the frequency with which a new approach-response was evoked; but for the test of new learning the positive stimulus was hidden from view, this measure being primarily one of determining the frequency with which a new (turning) response was made.

Results and discussion. The frequency with which each *S* ran to the positive goal-box for Trials 2-16 was determined. A critical ratio (*CR*)

TABLE I

FREQUENCY OF POSITIVE RESPONSE

Group	Test Condition	*M*	*CR*	*p*
I	Evocation	10.06	5.33	$<.01$
II	Learning	9.06	3.25	$<.01$

was then computed testing the mean observed frequency under each condition against the mean expected frequency by chance (7.5).

As can be seen in Table I, both groups ran to the positive goal-box significantly more often than would be expected by chance. Additionally, the frequency with which the *Ss* of Group I who were trained to run to the black box as the positive stimulus ran to that box during (evocation) test-trials ($M = 10.37$) was compared with the frequency with which white-positive-trained *Ss* ran to the black box during (evocation) test-trials ($M = 5.25$). A *t*-test indicated that the former ran to the black box significantly more frequently than did the latter ($p < 0.05$). Similarly, the *Ss* of Group II, who were trained black-positive, ran to the black box ($M = 8.25$) more frequently than the *Ss* who were trained white-positive ($M = 5.12$, $p < 0.01$). These composite findings are consistent and allow us to conclude that positive results were obtained using either the evocation or the new learning criteria. Hence, if, in an experiment of this kind, we find that a stimulus has acquired evocation-properties, we may infer that it has also acquired reinforcing properties, and vice versa.[4]

---

[3] M. R. D'Amato, Secondary reinforcement and magnitude of primary reinforcement *J. comp. physiol. Psychol.*, 48, 1955, 378-380.

[4] Following testing, all *Ss* were given 36 retraining trials after which the testing conditions for the two groups were reversed. Despite the negative transfer in switching testing techniques, Group II went to the positive stimulus under the evocation-condition significantly more often than chance ($p < 0.05$), while *p* was less than 0.06 for Group I when performing on the new learning test.

APPARATUS NOTE

A DEVICE FOR POSITIONING TWO SHUTTERS

In the course of our research on color-discrimination, we faced a somewhat difficult timing problem. It was necessary to expose a circular field of light with two successive flashes of different durations. The device pictured in Fig. 1 seems to do the job quite well.

Two self-cocking camera-shutters set at appropriate speeds are mounted on lens-

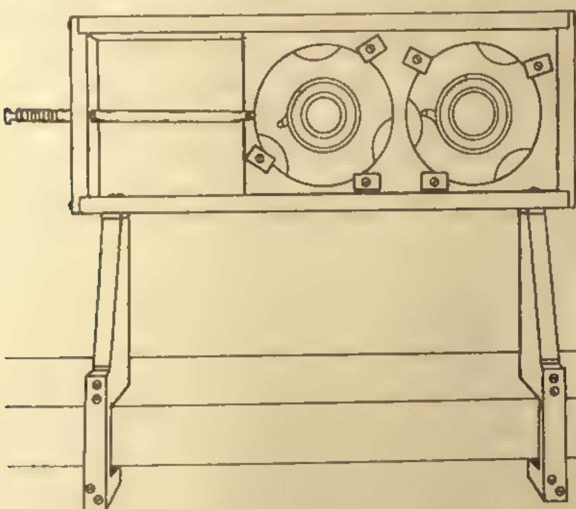

FIG. 1. DIAGRAM OF THE APPARATUS
In the position pictured, the shutter to the left would be directly in line with *O*'s eye. When the rod is pulled to the stop on the left, the right shutter would be directly in line with *O*'s eye.

boards. The lens-boards themselves are mounted on a plastic plate housed in a metal frame. *E* can position each of the shutters by means of the rod. When the rod is pushed to one stop, light comes through one shutter; when the rod is pulled back, light comes through the second shutter. With solenoids, it is possible to position the assembly from a remote location and thereby eliminate the manual operation of the rod. The shutters may be activated mechanically, electrically, or pneumatically.

This comparatively simple device can be constructed easily and inexpensively. We are not aware of a simpler workable device capable of programming two different exposure-durations.

U.S. Army Edgewood Arsenal

MICHAEL H. SIEGEL
EDWARD HOWARD

NOTES AND DISCUSSIONS

PHYSICAL AND PERCEIVED VISUAL EXTENT: A NOTE ON GILINSKY'S EQUATIONS

Gilinsky has proposed two equations relating physical and visually perceived spatial extent.[1] Her paper has been criticised on several grounds. Smith, Gruber and Rump have pointed out that the equations are contradicted by published data from a number of sources and Smith has pointed out a number of self-contradictions.[2] Without reference to their empirical validity however, Gilinsky's equations may be criticised on the grounds that the relationships that they specify between physical and perceived extent are non-invariant under transformations of the measurement-scales involved.

Gilinsky's first equation specifies the relationship between perceived distance and physical distance as,

$$d = AD/(A + D) \qquad [1]$$

where, $d =$ perceived distance; $A =$ the limit approached by perceived distance as physical distance approaches infinity; and $D =$ physical distance.

Discussion will be confined to this equation. The following considerations apply equally well, however, to Gilinsky's second equation, which relates perceived size to physical distance.

The status of physical distance in Equation [1] is, of course, clear. The scale-value for any distance is fixed by the choice of the measuring rod designated as unity and may be arbitrarily changed by change in this. In the case of *perceived* distance, however, the status of the variable is less clear. Gilinsky's discussion considers a scaling procedure by which a measure may be so attached to an object as to relate its judged distance to the judged length of a standard 'measuring rod' held at a standard distance. This standard would designate the unit of the scale of perceived spatial extent. Thus, if the standard were a meter-stick, the unit of perceived distance or size would be a 'perceived meter.'

---

[1] A. S. Gilinsky, Perceived size and distance in visual space, *Psychol. Rev.*, 58, 1951, 460-482.

[2] W. M. Smith, Gilinsky's theory of visual size and distance, *Psychol. Rev.*, 59, 1952, 239-243; H. E. Gruber, The size-distance paradox: A reply to Gilinsky, this JOURNAL, 69, 1956, 469-476; E. E. Rump, A note on "distance constancy in schizophrenic patients," *J. Ment. Sci.*, 107, 1961, 48-51.

In fact, the two experimental measures of perceived distance reported by Gilinsky do not employ precisely this technique. In one case, a simple fractionation-(bisection) technique was employed. In the other, a method involving the summation of successive perceptually equal increments in radial distance was employed. By any of these techniques, the scale-values obtained would not be unique. In each case, it would be possible to transform scale-values by an arbitrary change in unit. In the first case, this could be brought about by a change in the physical length of the standard measuring rod and in the latter two, by a change in the physical radial distance designated as unity.

The question arises as to the independence of the units employed for the perceived and physical distance scales. We will consider the case where the physical distance designated as unity on the physical distance-scale need not be the same as that designated as unity on the perceived distance-scale and where there are no restrictions on independent changes in unit for these two scales.

Suppose that Equation [1] held under a particular pair of units of physical and perceived distance. Then, designating values on these scales as D and d (and A) respectively, we could write,

$$d = AD/(A + D).$$

Suppose that the physical distance-scale were transformed by a change in unit so that D was replaced by kD. Then, this equation would become,

$$\begin{aligned} d' &= AkD/(A + kD) \\ &= dk(A + D)/(A + kD). \end{aligned} \qquad [2]$$

That is to say that (where $k \neq 1$), $d' \neq d$. This conclusion means that, given the value of A, the scale-value for perceived distance demanded by Gilinsky's theory changes with arbitrary changes in the unit of physical distance.

When the values of d and D are known, it would be possible to calculate the value of A demanded by Gilinsky's theory. Solving Equation [1] for A gives,

$$A = Dd/(D - d). \qquad [3]$$

Suppose, again, that this equation held for a particular pair of units of physical and perceived distance. Then solving for A for different paired values of d and D would always yield the same value of A. Suppose that the unit of physical distance were so changed that D was replaced in the equation by kD.
Then, Equation [3] would become,

$$\begin{aligned} A' &= kDd/(kD - d) \\ &= Ak(D - d)/(kd - d). \end{aligned} \qquad [4]$$

That is to say that (where $k \neq 1$), $A' \neq A$. Furthermore, since Equation [1] held for the initial physical distance-scale, we could substitute $AD/(A + D)$ for d in

Equation [4] and hence write,

$$A' = AkD/[D + (k - 1)A] \tag{5}$$

The factor multiplying A in this expression would change with changes in D and hence the value of A' would change with physical distance.

These conclusions provide something of a problem for Gilinsky's theory. Since A is a perceived distance, it is measured in the same units as d and the relationship between A and d should be invariant under transformations of the physical distance-scale. Yet it has been seen that given A, the value of d must change with changes in the unit of physical distance. Alternatively, given d, the value of A must vary with such changes in unit.

Gilinsky quotes a number of experiments in support of her theory. In her first experiment, she directly establishes a function relating d and D. She then employs Equation [3] above to obtain an estimate for A for every observed level of d. She reports the results for two subjects (*S*s). Consider those for *J.B.* reported in Gilinsky's Table I and Fig. 5. For this *S*, perceived and physical distance were measured in units of 'perceived meters' and meters, respectively. Approximately the same values of A were obtained at all levels of d and consequently the observed relationship between d and D was close to that predicted by Equation [1] (using the mean of the observed values of A). This procedure is a weak test of Gilinsky's theory, since it depends, not on confirming the determinate relation between d, A and D, but, rather, on observing whether the same value of A is obtained for different values of d. In that the different estimates of A were similar, however, the data do appear to confirm the theory. Suppose that the values of A had been calculated by employing feet instead of meters as the unit of physical distance. In this case, Gilinsky would in the first place have obtained different values of A. In the second place, the obtained values would have risen with d, thus apparently *contradicting* the theory.

The preceding discussion may be summarised as indicating that, for a particular scale of perceived distance, *if Gilinsky's equation held under one transformation of the physical distance-scale, then it could not hold under any other transformation.* Corresponding conclusions would apply to transformations of the perceived distance-scale.

The above considerations apply to the case in which no restrictions are placed on independent changes in the units of perceived and physical distance. Smith, however, has suggested that Gilinsky intends that perceived and physical distance in her equations should be measured in the

same units.[3] (That is, the physical distance designated as unity on the perceived-distance scale should be the unit of physical measurement.) Support for Smith's suggestion is to be found in the fact that Gilinsky does, in her examples, always use the same units. The question might be raised as to whether this expedient would make her equations tenable.

Suppose that Equation [1] held for a particular pair of units of perceived and physical distance. The equation would still hold under transformations of the physical distance-scale, provided that these were always accompanied by corresponding transformations of the perceived distance-scale. Thus, if in Equation [3], values of d were replaced by kd and values of D, by kD we would have,

$$\begin{aligned} kA &= k^2Dd/k(D-d) \\ &= kDd/(D-d), \end{aligned} \qquad [6]$$

and hence Gilinsky's equation would still hold. For example, if the equation happened to hold for yards and perceived yards, it would also hold for feet and '⅓ perceived yards' (*i.e.* values on a scale of perceived distance in which the unit was the physical distance judged to be ⅓ of a physical yard). *Except, however, where perceived and physical distance were linearly related (and hence according to Gilinsky, $A = \infty$), the ⅓ perceived yard would not be a perceived foot.* That is, the expedient of insisting that the physical and perceived distance-scales have the same unit would not normally make Gilinsky's equations tenable.

It was stated in the last paragraph that Gilinsky's equations could hold, given certain restrictions on independent changes of unit for the two distance scales. In this case, her theory must be reformulated as the assertion that, *given a particular transformation of the physical distance scale (or the perceived distance-scale), there is one transformation of the perceived (physical) distance-scale such that Equation [1] holds.* This was not made clear in Gilinsky's article. In addition, it might be pointed out that (without raising the general question of dimensional homogeneity), it is not the practice in the physical sciences to allow laws which hold only for particular units of measurement (through Luce's weak restrictions on admissible laws would lead to the acceptance of such).[4]

Testing the theory in this form would not be quite straightforward. If A were directly determined, its truth could be easily established by selecting arbitrary units for d and D, determining pairs of values of d and D,

---

[3] Smith, *op. cit.*, 241.
[4] H. D. Luce, On the possible psychophysical laws, *Psychol. Rev.*, 66, 1959, 81-95.

and then finding in each case a transformation of D (or d) such that Equation [1] held. Gilinsky's theory requires that this transformation be the same in each case. Gilinsky, however, suggests no method by which A might be directly determined. Instead, the tests that she suggests are all based on establishing the consistency of the data in terms of her theory. Her first experiment (described above) relied upon showing that the same value of A was obtained by applying Equation [3] to different pairs of values of d and D. We have seen, however, that the success of such a method would depend on the units that happen to have been selected for these two variables. If this technique were modified, it would be necessary arbitrarily to select units for d and D. It would then be necessary to show that there was one transformation of D (or d) such that the values of A given by Equation [3] were constant over different levels of d (or D). The only technique available for this task would appear to be trial and error.

Gilinsky's equations were not obtained empirically but were derived rationally as a result of applying simplifying assumptions to Luneberg's theory of visual space.[5] The question therefore arises as to how the observed non-invariances occurred. Reference to Gilinsky's derivation makes this clear.[6] She has, at a number of points, made assumptions which are themselves non-invariant under scale-transformations. In particular, she has employed the substitution $d = a/\gamma + \mu$ (where $a =$ interpupillary distance, $\gamma =$ angle of binocular convergence and $\mu =$ the threshold of perceived convergence). As in the case of Gilinsky's final equations, this equality is non-invariant under independent changes in the unit of either physical or perceived distance. There does not appear to be any way out of this difficulty.

Australian National University M. L. Cook

EDWARD WHEELER SCRIPTURE: 1864-1945

Edward Wheeler Scripture, one of the younger pioneers of the "new psychology" in America and in fact the man who brought the phrase *new psychology* into use as indicating the non-philosophical experimental psychology that Fechner, Helmholtz, and Wündt introduced about 1860,

[5] R. K. Luneberg, *Mathematical Analysis of Binocular Vision*, 1947, 37.
[6] Gilinsky, *op. cit.*, 476. In addition to the derivation from Luneberg, Gilinsky offers derivations of the same equations from the laws of geometrical perspective and the size-distance invariance hypothesis respectively. In neither case, however, are these derivations in any way rigorous.

was born in Mason, New Hampshire, on May 21, 1864, and died near Bristol, England, on July 31, 1945, at the age of 81.

This note is published to establish the correct date of Scripture's death in 1945, an event which was overlooked when it occurred and has only recently been correctly printed. Scripture's role as one of the young pioneers in American psychology is well known.[1] He was certainly one of the sixteen or more psychologists present in G. Stanley Hall's study at Clark University at the founding of the American Psychological Association on July 8, 1892.[2] Not much is known as to why he left Yale in 1903 after having been Professor of Experimental Psychology and Director of the Psychological Laboratory there, but he lists without dates in *Who's Who* (British) and *Wer Ist's?* Associate in Psychiatry and Director of the Research Laboratory in Neurology at Columbia University, Investigator in the Carnegie Institution of Washington, Lecturer on Experimental Phonetics at the University of Marburg, Honorary Lecturer on Phonetics at King's College, London, and finally Professor of Experimental Phonetics at the University of Vienna. He went abroad in 1912, just before the First World War, carrying his interest in phonetics with him. His first entry in *Who's Who* gives for him a London address in 1917. He continued throughout the years to record British addresses, but added an address in Vienna from 1925 to 1935, which is presumably the period of his professorship at Vienna. After that *Who's Who* shows various British locations, until he settled down at Henleaze, near Bristol, where he died on July 31, 1945.

The confusion in respect of the date of Scripture's death was caused by the fact that, a failing older man, he dropped out of professional attention and his death received no professional notice, and further by the fact that the English *Who's Who* continued his sketch until the 1959 volume (a result of a procedure which would not have happened with *Who's Who in America*). In the late 1940's I found an excuse to write to Scripture at Henleaze in order to see if he were still living, and I received no reply. Then in 1955, W. L. Bryan died and the necrologist in this JOURNAL stated that Bryan was the last survivor of those present on July 8, 1892, at the founding of the APA in Stanley Hall's study. As editor, I, questioning this statement because I thought Scripture was probably still living, wrote to the Secretary of the British Psychological Society and to the American Embassy in London to see if I could find out the

---

[1] E. G. Boring, *A History of Experimental Psychology*, 2nd ed., 1950, 527 f.

[2] Wayne Dennis and E. G. Boring, The founding of the APA, *Amer. Psychologist*, 7, 1952, 95-97.

truth. My query to the Society resulted in a Bristol psychologist's going to the address at Henleaze to be told that Scripture had died many years before. My query to the Embassy brought the report that the Alien Office had recorded his death as August 5, 1945, a six-day error which unfortunately was recorded by me in a footnote to the Bryan necrology.[3] How hard it is to be right! (Nor did this discovery make Bryan the last survivor of Hall's founding group of July 8, 1892, for Lightner Witmer, who was there, did not die until July 1956, eight months after Bryan.)

For nine years this obscure footnote with its six-day error seems to have been the only published record of Scripture's death. I tried to get someone to write a biographical sketch of the late Scripture, but no pronetic psychologist of my acquaintance found this undertaking more interesting than what were already his most pressing commitments. Then in January 1965 the house organ of the American Speech and Hearing Association published a short biography of Scripture with the true date of his death in it, taken, its author said, from the death certificate.[4] A copy of this surprising sketch was sent me by Dr. Ira J. Hirsh of the Central Institute for the Deaf in St. Louis. Having already been misinformed by the American Embassy in London, I asked Professor C. A. Mace of the University of London to check for me, and he sent me a photostat of the death certificate which he obtained from the General Register Office at Somerset House in London. (Even it has an error, for it gives Scripture's age as 82, whereas arithmetic makes it 81. There is no ambiguity about the date of birth.)

You would think that there would be plenty of information about Scripture's later life, for Carl Murchison printed his bibliography in 1932 and his autobiography in 1936. The bibliography gives 216 references from 1891 to 1932, to which may be added 11 in 1933-1935 from the G. K. Hall volumes.[5] In the period 1926-1932, almost all of the articles—44 of them—are in German. All the rest of the 225, save for one in French, are in English. The alleged "autobiography" contains only two dates and is thus almost entirely concerned with intellectual reminiscences about people, ideas, and research. There really is no *vita* at all. Scripture may have had the compulsions of an egoist, but he was anything but an

---

[3] D. G. Ellson, William Lowe Bryan: 1866-1955, this JOURNAL, 69, 1956, 325-327.

[4] M. F. Berry, Historical vignettes of leadership in speech and hearing; I. Edward Wheeler Scripture: 1864-1945; *Asha*, 7, Jan. 1965, 8 f.

[5] Carl Murchison, ed., *The Psychological Register*, 3, 1932, 426-440; Columbia University's Psychology Library, *Author Index to Psychological Index and Psychological Abstracts: 1894-1958*, G. K. Hall, publ., 1960, V, 3446 f.

exhibitionist. At the end of his autobiography, he appends this afterthought:

"I notice a paucity of personal details in my account. I have forgotten most of them and am not interested in the rest; I do not think the reader would be interested either. In order to be dated and placed I have to state that I was born in 1864 in a village in New Hampshire, U.S.A."[6]

Dated and placed! Almost, he failed to be. Perhaps that was his wish.

Harvard University — EDWIN G. BORING

ANCHOR-EFFECTS AND THE SEMANTIC DIFFERENTIAL

When he introduced the semantic differential, Osgood maintained that the instrument was relatively independent of set or anchor-effects and cited an unpublished study by Aiken in support of this view.[1] During the past three decades, however, many studies have shown that judgments of size, distance, brightness, social status, and, more recently, of clinical diagnosis are susceptible to contrast- and assimilation-effects.[2] Since the semantic differential is widely used and the same subject (*S*) is asked to rate a large number of concepts, it seemed important to learn if Osgood's assumption regarding the lack of anchor-effects is correct. It should be noted that we make no specific hypothesis as to whether contrast or assimilation will be evidenced. Conceivably, a test-item like *politician* could be assimilated into a perceptual unit *thief-liar-fool;* on the other hand, since each item is rated separately, it is possible that a politician might appear in a more favorable light in the company of thieves, liars, and fools.

The plan of the study called for each *S* to rate three set-inducing words and then a test-word. The three set-inducing words were intended to be homogeneous in some certain characteristic. For example *thief, liar,* and *fool* were all low in value, while *janitor, garbage collector,* and *farm laborer* were all low in status. Other *S*s drawn from the same population

---

[6] E. W. Scripture, in Carl Murchison (ed.), *A History of Psychology in Autobiography,* 3, 1936, 231-261.

[1] C. E. Osgood, G. J. Suci, and Percy Tannenbaum, *The Measurement of Meaning,* 1957, 84-85.

[2] James Bieri, B. A. Orcutt, and Robin Leaman, Anchoring effects in sequential clinical judgments, *J. abnorm. soc. Psychol.,* 67, 1963, 616-623; D. T. Campbell and Nan Lewis, The effects of assimilation and contrast in judgments of clinical material, this JOURNAL, 70, 1957, 347-360; O. J. Harvey and D. T. Campbell, Judgments of weight as affected by adaptation range, adaptation duration, magnitude of unlabeled anchor, and judgmental language, *J. exp. Psychol.,* 65, 1964, 12-21; E. G. Wever and K. E. Zener, The method of absolute judgments in psychophysics, *Psychol. Rev.,* 35, 1928, 466-493.

would rate three items intended to establish the opposite set—for example, *stateman, scholar,* and *scientist*—followed by the test-item. The difference between the ratings of the test-item should reflect the operation of the effect of set.

The same general procedure was followed on five different occasions using five different classes of students in psychology. The same *S*s were never tested twice. The positive and the negative forms were distributed randomly within each class. The first two pages of each booklet contained standard instructions for the semantic differential.[8] Then followed four pages, each containing the same nine scales, three representing each of the three major dimensions of meaning—value, activity, and potency. Some *S*s rated three negative concepts and then a test-item, while other *S*s in the same class received three positive items followed by the test-item. The five test-items were: *expert, politician, carpenter, family,* and *winter.* For each test-item, there was an appropriate series of positive and negative set-inducing items, *e.g.* winter was preceded either by *hate, steal,* and *devil* or by *faith, peace,* and *sunlight.*

As 5 separate test-items were rated along each of 9 scales, 45 different comparisons were possible. Only five of the differences were significant beyond the 5% level by the *t*-test. Furthermore, the largest of the differences was not more than one scale-unit on the seven-unit scale. Most of the other differences were much smaller than this.

These results show that the semantic differential is markedly resistant to the operation of anchor-effects. Since the conditions of the present series of studies were designed to maximize such effects by bunching positive or negative items, the user of the semantic differential can have even more confidence in his results if his items are arranged randomly. The present study indicates that even if some positive or negative items should cluster together, they should not significantly affect the direction or magnitude of the ratings. It also suggests that the judgmental process involved in the semantic differential differs from that involved in many other procedures, in that the stimuli rated appear to be treated discretely rather than as part of a series to be judged using a single external yardstick. Some further investigation of the classes of judgment that are *not* susceptible to anchor-effects would appear to be warranted.

University of California, Davis ROBERT SOMMER

[8] Osgood, *op. cit., passim.*

COMMENTS ON "SUBJECTIVE COLOR: A NEW METHOD OF PRODUCING THE PHENOMENON"

Knehr and Lorenz have recently reported that when a black and white patterned disk is rotated at the same frequency as the flickering of a fluorescent light source, a well-saturated pattern of orange and blue can be seen.[1] The authors assume this effect to be a new method of producing the well-known subjective color phenomenon.[2]

There is, however, an alternative to this explanation. A peculiarity of most fluorescent light is that each pulse or "flicker" contains a brief but distinctly orange and a complementary blue phase. In a 60 cycle A.C. circuit these phases follow one another so quickly that they fuse, and the resulting light appears white. Apparently, the effect obtained by Knehr and Lorenz occurs because of the separation of these orange and blue phases that results when the disk is rotated in synchrony with the flickering of the light. For the simplest pattern they report using, the half-white-half-black disk, this separation comes about as follows: the white portion of the disk is first illuminated by the blue phase, then, after rotating 180°, it is lighted by the orange phase, then it rotates 180° farther to catch the next blue phase, and so on. The result is that the disk appears to be stationary with rough halves of relatively well-saturated orange and blue. These colors may also be observed directly by viewing the fluorescent light through a disk having a narrow radial slot when this disk is rotated at 60 r.p.s. Satisfactory color photographs of these hues have been obtained using both methods.

While these procedures provide an interesting demonstration that the apparent white of fluorescent light contains distinct orange and blue phases alternating above the *CFF*, they should not properly be considered a new method of producing the subjective color phenomenon.

University of Wisconsin RICHARD E. GORANSON

THE 1965 MEETING OF THE AMERICAN PHILOSOPHICAL SOCIETY

The major papers of the meeting were devoted to a commemoration of the publication in 1865 of Gregor Mendel's *Pioneer Experiments in Genet-*

[1] C. A. Knehr and R. J. Lorenz, Subjective color: A new method of producing the phenomenon, *J. Psychol.*, 58, 1964, 353-356.

[2] For an historical review of this phenemenon, which is continually being rediscovered, see M. B. Erb and K. M. Dallenbach, Subjective colors from line patterns, this JOURNAL, 52, 1939, 227-241.

ics. In a number of these papers, reference was made to modern views on the genetics of behavior.

There were a number of other items of interest to psychologists in the papers. It was noted that the inheritance of color-blindness was studied at an early period. Goddard's genetic theory of feeble-mindedness was also referred to, as were the studies of Freeman and others on the psychological characteristics of twins reared apart. In a paper on Rittenhouse, reference was made to the fact that this distinguished 18th century American scientist had written on visual illusions and this work had been referred to by Helmholtz and Brewster.

The psychologists in attendance were F. A. Beach, E. G. Boring, L. Carmichael, W. Köhler, W. R. Miles, and C. Pfaffman. Heinrich Klüver of the University of Chicago was elected to membership as a psychologist, and Hallowell Davis, the distinguished student of audition, was elected as a physiologist.

The award of the Karl Spencer Lashley Prize (a diploma and $2000) to Dr. Giuseppe Moruzzi, Head of the Institute of Physiology of the University of Pisa, Italy, was announced. The prize will be presented at a later time when Professor Moruzzi can come to Philadelphia.

National Geographic Society LEONARD CARMICHAEL

THE 1965 MEETING OF THE NATIONAL ACADEMY OF SCIENCES

There were a number of items on the program of interest to psychologists. A symposium was held on the "Mechanisms of Color Vision" under the chairmanship of Clarence H. Graham of Columbia University. The speakers were Deane B. Judd, National Bureau of Standards: "Fundamental Studies of Color Vision from 1860 to 1960;" Edward F. MacNichol, Jr., Johns Hopkins University: "Retinal Processing of Visual Data;" David H. Hubel, Harvard Medical School: "Effects of Varying Stimulus Size and Color on Single Lateral-Geniculate Cells in Rhesus Monkeys;" and George Wald, Harvard University: "The Retinal Basis of Human Color Vision and Color Blindness."

Harry F. Harlow of the University of Wisconsin gave a paper on the "Effects of Total Social Isolation on Macaque Monkey Behavior," and W. K. Estes of Stanford University on "A Technique for Assessing Variability of Perceptual Span." Leonard Carmichael of the National Geographic Society showed a film by Jane Goodall and the Baron Hugo van Lawick on a field study of wild chimpanzees.

Neal E. Miller was elected Chairman of the Section on Psychology of the Academy and Wendell R. Garner of Johns Hopkins University was elected to membership in the Academy.

The psychologists in attendance were F. A. Beach, L. Carmichael, W. K. Estes, C. H. Graham, H. F. Harlow, W. R. Miles, W. D. Neff and Carl Pfaffman.

National Geographic Society LEONARD CARMICHAEL

SIXTY-FIRST ANNUAL MEETING OF THE SOCIETY OF EXPERIMENTAL PSYCHOLOGISTS

The sixty-first annual meeting of the Society of Experimental Psychologists was held at The Ohio State University on April 2 and 3, 1965. Delos D. Wickens, Chairman of the Society for 1964-65, served as Chairman of the meeting.

Members and fellows present were: Beach, J. S. Brown, Deese, Duncan, Egan, Fitts, Gagné, Garner, Geldard, Grant, Helson, Irwin, Kappauf, Kendler, Kennedy, Kimble, Lehrman, Mackworth, Melton, Meyer, Miller, Mueller, Neff, Olds, Pfaffman, Postman, Riggs, K. U. Smith, Solomon, Underwood, Volkmann, and Wickens.

Abram Amsel, Gordon H. Bower, David Krech, and Roger N. Shepard were elected to membership in the Society, bringing the total to 67 members and 27 fellows.

Reports of research in progress were presented in scientific sessions during the afternoon of April 2 and on April 3.

The Warren Medal for 1965 was awarded to William C. Young "in recognition of his pioneering investigations of hormonal factors controlling reproductive behavior of mammals, from rodents to primates, and his singular success in combining the techniques of experimental psychology with those of endocrinology to increase our understanding of the physiology of motivation."

The Society accepted the invitation of Duke University to meet in Durham, North Carolina in 1966. G. A. Kimble was elected Chairman for 1965-66.

Princeton University JOHN L. KENNEDY

THIRTY-SIXTH ANNUAL MEETING OF THE EASTERN PSYCHOLOGICAL ASSOCIATION

The Eastern Psychological Association held its thirty-sixth annual meeting April 22-24, 1965 at the Chalfonte-Haddon Hall in Atlantic City, N.J.

Total registration at the meeting was 2,651. Of this number 1,179 were members of EPA, 349 were new members who joined the Association at the meeting, and 1,123 were guests. The present membership of EPA totals 4,084.

Lewis W. Field was chairman of local arrangements, assisted by Murray Benimoff, Peter Finley, Robert K. McKelvey, Edmund Shimberg, and Carter Zeleznik. The Program Committee comprised Bernard Saper, Chairman, Leonard W. Ferguson and L. Starling Reid. The Committee scheduled 376 papers (presented in 74 sessions), 11 symposia, 4 special sessions, 1 film session, 7 special group meetings, and 3 invited addresses.

Of the 74 paper-reading sessions, 50 were devoted to topics in physiological psychology, animal behavior, and the experimental study of perception, learning, and motivation. Some of the more specialized titles of sessions in these areas were as follows: psychopharmacology, brain lesions, brain stimulation, motivation-physiological, animal motivation, operant behavior, generalized behavior, punishment, avoidance behavior, reinforcement, audition, vision, visual-motor behavior, information processing, cognitive processes, verbal behavior, memory, serial learning, associative transfer, and paired-associate learning.

Three paper sessions were concerned with developmental and child behavior. Social psychology was covered by eight sessions, clinical research by four sessions, and personality-studies by three sessions. The remaining six sessions related to measurement, educational psychology, industrial psychology, and human factors.

Commemorative symposia were held in honor of Harold Schlosberg and Heinz Werner. Brief titles of other symposia were "The investigation of personal time," "Experimental psychology of reading," "The young business manager," "Perception without awareness," "Conditioning techniques in physiological studies," "Cognitive functioning of elderly people," "Cognitive patterning of complex stimuli," "Psychotherapy as behavior change," and "Theory formation in psychology."

J. P. Scott gave an invited address, "The control of social behavior." By special invitation, M. B. Shapiro of Maudsley Hospital, London, England spoke on "The single case in clinical psychological research." Psi Chi had James J. Gibson as an invited speaker. Bartley G. Hoebel, winner of the AIR Creative Talents Award, addressed a session on "Some studies of feeding, drinking, anl self-stimulation." David C. McClelland delivered the Annual Presidential Address, "Developing achievement motivation."

Among business transactions, Eliot Stellar was announced as the Presi-

dent-elect and Charles N. Cofer and James Deese as new members of the Board of Directors. Arthur A. Witkin was elected to a three-year term as Treasurer. The Board of Directors voted to recommend a change in the by-laws which would assign the function of disbursing funds to the Treasurer and leave the collection and assignment of funds to the Executive Secretary. Other actions of the Board related to the allocation of $3,500 from surplus funds to underwrite travel fellowships of young EPA psychologists for attending the 1966 International Congress of Psychology in Moscow and the consignment of $500 to an award to be made by the President-elect. Jeffrey P. Goldner of Queens College received this year's President's Award for having the outstanding paper among undergraduate students who reported at the meeting on their individual research.

The 1966 meeting of the Association will be held at the Statler-Hilton Hotel in New York City, April 14-16.

Queens College MARVIN A. IVERSON

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Problems in measuring change. Edited by CHESTER W. HARRIS. Madison, University of Wisconsin Press, 1963. Pp. x, 259. $7.50.

The aim of this set of papers is a psychometric view of the measurement of change.[1]

Since the imposing array of 12 papers by 14 well known authors cannot be discussed adequately even in an extended review, only those topics of special interest to the experimental psychologist will be considered.

Probably the chief news of importance to experimental psychologists is contained in the Gaito and Wiley paper on the univariate analysis of variance and R. D. Bock's paper on multivariate analysis of variance. "For univariate analysis procedures to be correct, the covariance matrix should have equal variances in the principal diagonal and . . . *constant co-variance* elsewhere." (Gaito and Wiley, p. 72). Bock (p. 86) points out the difficulties that arise from this constant-correlation requirement. "In (agricultural) field trials there appear to be good . . . reasons for believing that the variation due to sampling plots will be of a form amenable to univariate analysis of variance." But repeated measurements on the same set of organisms generally give rise to correlations which decrease away from the main diagonal, *i.e.* the matrix resembles the Guttman "simplex" type. No useful exact univariate analysis for the repeated measurement design appears feasible for this type of correlation matrix. In general multivariate analysis of variance is necessary (although Geisser and Greenhouse, following the lead of G. E. P. Box, have shown that a conservative univariate test is possible).

Recognition of the essentially multivariate nature of the repeated-measurements design is rather rare among psychometricians. Most statistical psychologists make the mistake of equating the repeated-measurement design, which uses the same organisms, to the split-plot agricultural design which used separate blocks of earth. Such inappropriate univariate analysis of variance procedures are found in most statistical psychology texts (*e.g.* A. L. Edwards, *Experimental Design in Psychological Research,* 1960; E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education,* 1953; and Q. McNemar, *Psychological Statistics,* 1962).

For most of us, D. T. Campbell's chapter on experiments involving time-series will raise questions far more important than the problems of statistical analysis. After all, psychologists don't really need much in the way of statistical analysis if they are willing to repeat their experiments, but if the experimental design puts systematic bias into the results, no amount of repetition will help. Campbell lists twelve threats to experimental validity; in my experience the most commonly en-

[1] The preparation of this review was partially supported by National Institute of Health Grant MY-5504. The opinions and assertions contained herein are the private ones of the writers.

countered difficulties in the repeated measures within Ss' design are *treatment carryover* and *test carryover*.

'Treatment carryover' can occur whenever two or more treatments are applied to the same subject. There may simply be treatment carryover to the following treatment, in which case the crossover designs of the agricultural statisticians are adequate. But there may be a residual interaction, a potentiation or inhibition of the effects of subsequent treatment, (vaccination is a dramatic example of an inhibiting interaction) in which it is usually impossible to ascertain the unbiased effect of the separate treatments (unless special *rotation* designs are used).

'Test carryover' occurs when the measurement of a subject has a constant residual effect on a subsequent treatment or when the measurement effect interacts with the subsequent treatment effect. It also occurs as a quasi-Heisenberg effect, *i.e.* it may alter subsequent measurements by altering the state of the organism. There is an interesting sociological problem here. In contrast with their attitude to treatment carryover, statisticians (except for a handful) are profoundly uninterested in test carryover. Psychometricians recognize the test-carryover problem but generally refuse to deal with it, as can be seen by this typical excerpt from the book under review. "We choose to ignore learning or other temporal effects, such as fatigue . . ." (Webster and Bereiter, p. 43). To use Jane Loevinger's phrase, they are only interested in the "Theory of the Single Test."

But social psychologists are quite aware of test-carryover effects and have even devised special .experimental designs to assess possible interactions with subsequent treatments. Studies using these designs have shown that attitude-tests can interact significantly with subsequent influenced treatments.

Lord, in his usual light-hearted style, discusses two psychometric paradoxes. The first is the well-known "regression to the mean." Let X and Y be two successive measures on the same individual where all the moments of X (mean, variance, skewness, etc.) are the same as for Y and the correlation between X and Y is less than unity. Then for subjects with very high values of the initial measure, X, it is virtually certain that the second measure, Y, will be less than X. Similarly, for very low values of X, Y will be greater than X. It is easy to see how this can occur when the true scores are equal, $T_x = T_y$, so that the observed X and Y differ only because of measurement error. But what if there is no measurement error? Can we still have "regression to the mean?"

"Dynamic equilibrium" is the phrase Lord uses for the case where there is a non-zero true change, $T_g = T_y - T_x$, but the group of subjects maintains the same score distribution. Lord shows that for the distribution to be stationary, the true change must have a negative correlation with the true initial measure,

$r_{t_g t_x} = -S_{t_g}/2S_{t_x}$.

Do dynamic equilibria exist? Yes, whenever there are antagonistic neura mechanisms exerting homeostatic effects, (as for example in heart rate and respiration rate) stimulating a subject who has a low resting level generally leads to an increase in rate-stimulating a subject with a high initial value generally leads to a decrease in rate. This homeostatic effect has been given the name "Law of Initial Values" by J. Wilder and is well known to psychophysiologists through the writings of J. I. Lacey.

All this has led Benjamin, Manning, and DuBois and others to assert that the deviation from the regression line, $Z = Y\text{-}a\text{-}bX$ (the residual change-score), is the 'real' measure of change and the usual difference score, $D = Y\text{-}X$, is a false measure of change. Lord resists this conclusion: he contends that it is generally more useful and convenient in comparing treatments to have a change-score that is independent of initial level, but the difference score or some other measure of change may be most appropriate for the question being asked.

The second of Lord's paradoxes concerns sampling fluctuations and the analysis of covariance (*anacova*). Consider the case where individuals are assigned *strictly at random* to treatment Groups M and N. A pre-treatment score X and a post-treatment score Y are obtained for each subject. The null hypothesis is correct, there is no treatment effect. Lord shows that if the measures are infallible, and the population is in a state of dynamic equilibrium, *anacova* can fully compensate for sampling fluctuations in the covariate X. However, suppose that (a) the measures are fallible, (b) the population is in a state of "time-independent" equilibrium (the usual assumption of the psychometrician) and (c) the group averages on X differ significantly. Then *anacova* is biased, and the analysis of differences will compensate for sampling fluctuations. So using a measure of change that is not corrected for pre-treatment difference may be better than using the residual change score!

One way of resolving the paradox is to remember that in those cases where the two groups have significantly different pre-treatment means, we are looking at only a small part of the set of experiments that would be generated by random assignment of subjects into two equal groups. Presumably over the entire set of assignments, the *anacova* would give the correct answer more often than the analysis of differences. Furthermore, as the number of subjects increase, $\overline{X}_m$ must approach equality with $\overline{X}_n$.

The statistician would say that the *anacova* has greatest power over all possible experiments. He could recommend using a 95% confidence interval in conjunction with *anacova* and be very confident that when the null hypothesis was true, 95% of the research workers would accept the null hypothesis.

But no confidence man can help the 5% who are making the error of rejecting the null hypothesis when it is true, in particular those experimenters who find that their randomized groups differ significantly on the pretreatment variable. *E* himself can, however, practice preventive medicine. If it is possible to obtain pre-treatment scores before the assignment to treatments, then randomized balanced blocks can be used. The procedure would be (a) to rank all *S*s on the pre-treatment score; (b) strictly at random, to assign the first block of *k* *S*s to the *k* treatments; then the next block of *k* *S*s to the *k* treatments, etc. In this way there is no chance that the *k* pre-treatment averages can differ significantly from one another. Such randomized matched assignments are astonishingly rare in psychological experiments. Where it is possible to take two or more measurements of the control variable, Lord (1960) has given a large-sample modification of *anacova* that is appropriate for fallible control measures.

What about the case where we are comparing groups that were not randomized by *E*? For example, schizophrenics vs. depressives, or first-born vs. second-born, and so on. As Lord states (p. 38) "If . . . randomized assignment is impossible,

then there is often no way to determine what is the appropriate adjustment to be made for initial differences between groups, and hence often no way to show convincingly by statistical manipulations that one treatment is better than another." *Anacova* is appropriate only when the population between-group regression is the same as the population within-group regression. If the *Ss* are not assigned randomly to the treatment-groups there is no guarantee that this is true. But surprisingly, many statistical psychologists (*e.g.* McNemar, *Psychological Statistics,* 1962) will argue, in the face of logic and the authority of the inventor, R .A. Fisher, that *anacova* is the analysis of choice when the experiment is performed on preëxisting non-randomly selected groups with significant group differences on the control measure. Possibly there are aspects of the *anacova* ritual that serve non-experimental needs.

And so, as we leave this 'Psychometric Land of Change,' faint murmurs pursue us "But what is the *real* measure of change?" The answer, of course, is that there is no one best way of measuring change.

Neuropsychiatric Research Unit — ARDIE LUBIN

Motivation as Related to Personality. By DOROTHY RETHLINGSHAFER. New York, McGraw-Hill Book Co., 1963. Pp. ix, 388. $7.95.

Motivation: Theory and Research. By C. N. COFER and M. H. APPLEY. New York, John Wiley and Sons, Inc. 1964. Pp. 958. $12.50.

It is often difficult to understand why books dealing with a particular topic suddenly appear in such numbers that it is safe to speak of a deluge. Such has been the case for "motivation," which has yielded about a dozen volumes within the brief span of about five years. This is a quite remarkable outpouring in a subject matter which for almost 29 yrs. was represented in American psychology by only one advanced textbook, *viz.,* P. T. Young's *Motivation of Behavior.* What is even more remarkable, in this reviewer's judgment, is how a subject which has been at the core of some of psychology's most interesting and important controversies could for so long lie dormant as a topic for advanced texts. Surely it could not be because Young had provided so thorough and satisfactory an account that there was no need for other interpretations; and, surely it could not be that there were no new ideas or new findings which warranted attention. No! One has to look at the question in a different light. True, there were no other books entitled *Motivation* during each of the pre and post World War II decades, but that is not to say that the topic of motivation was being ignored. It had, to some extent, gone underground during the '20s and '30s: instinct was terribly disreputable; consciousness was under objectivistic and behavioristic attacks, as were also theoretical and hypothetical explanatory mechanisms; and, psychoanalysis, whose influence rose to a peak during the '30s, became particularly identified with motivation and tarnished it with a metaphorical language that seemed to continue an old, and rejected, demonic view of mind. On the other hand, there were more positive factors at work: Tolman and Lewin, in rather different ways, to be sure, emphasized the purposive character of behavior; the physiology of nutritive and appetitive behavior came under concerted attack; the terms "drive" and "dynamism" emerged as substitutes for motivation; Hull's incentive motivation, drive reduction, and the multiplicative sources of $S^{E}R$ could be regarded as a theory of motivation; and, if

one sought a general account of motivation it could be found in books on personality and social psychology.

Why, then, should there now be a resurgency of interest in motivation as a general topic instead of a continuation of the relatively independent approaches with which psychology had become comfortable? Despite their differences, the two books under review give us some clues. First, it seems clear that functionalism has completely won the field; motivational phenomena are functional processes which assure the maintainance and the integrity of organisms. Cofer and Appley explicitly favor an equilibratory model which, it should be noted, is considerably more general than the usual homeostatic formulations; Rethlingshafer, while explicitly rejecting what she also describes as an equilibrium-theory, in fact proposes an approach that is similar to Cofer and Appley's. The latter authors urge a two-fold version of motivation to serve as the core of motivational theory: *Sensitization-invigoration* and *anticipation-invigoration.* The key term, *invigoration,* is not unlike the one which Rethlingshafer favors, vis. activation. In any case, the homeostatic, drive-type theory which relies upon some kind of deficit state to get everything going, is clearly rejected as a general model. The consequence is that organisms are seen as far richer and more powerful in their resources for survival and for mastering themselves and the task-setting environment. Instinctive behavior, appetitive and avoidance activities, hedonic processes, personal style or pace, blocking or conflict, the Self, social interaction, and the "new" complex of cognitive sources of arousal, though all differing among themselves in their details, come together happily once it is not necessary to find a unique deficiency condition for each of them.

A second factor that emerges from these two volumes is the enormous amount of empirical research which the past thirty years has produced, most of which, indeed, is actually post WW II. From the few, patchy, poorly executed, and ill-defined investigations prior to WW II on which ideas about motivation had to be based, psychology has moved into an embarrassing wealth of studies covering both complex and simple aspects of motivational behavior. The extensive studies of McClelland and his colleagues on *nAch;* of Sears and his group on aggression and dependency; of the Iowa group initiated by the application of the Taylor *MAS;* learned drives as Miller and others have studied them; motive conditions in relation to vigilance, signal detection, and perceptual organization; direct neural stimulation by electrical and chemical means; the interaction of incentive learning, deficit, and performance; the existence of neural receptors, monitors, and activators in relation to deficits; these and many other totally new lines of inquiry and fact are elaborated by Rethlingshafer and Cofer and Appley in their efforts to bring order into the field.

Further, it is clear that there has been more than just a quantitative increase in facts and experiments. Things have begun to fall into an hierarchical model so that, for example, the notion of invigoration or activation takes on a much more constrained and technical sense than the old notion of tension ever had a chance to, and, by virtue of the introduction of some kind of feedback notion, invigoration begins to take on some sensible attributes as an operating mechanism in behavior. In any case, the monolithic, master-motive type of thinking is all but gone; in its stead we find a complex of local and boundary conditions dealing with species attributes, learning, bodily states, local dispositions, and cultural and contingent

setting factors which converge on restricted behavioral measures such as response probabilities, behavior amplitude, or scale value.

A third feature of these two volumes is, for me at least, the most heartening of all. Whatever one's bias on a particular issue may be, I think that psychologists in general ought to rejoice in the theoretical freedom and flexibility which these many pages display. There are now extensive enough subsets of data and empirical laws in several fields that investigators can and do hypothecate suitable substantive and theoretical mechanisms. These are not the webspun fantasms of psychologists who have all of behavior and little data to deal with, but the free imagination of people who have lots of healthy data; an opportunity to generate relevant experiments; no father-figure peering over their shoulder whose dicta they must heed; and, a sublime indifference to warnings that certain things are now still too difficult to study or "really aren't psychology." How long this will last I don't know but psychologists ought to enjoy it as long as it does.

Now for a few comments about the differences between the two books. Cofer and Appley's is considerably larger, though the communality in topics and references is striking; to be sure, Rethlingshafer subtitles her book "as related to personality," but it is in fact not all that different in content from the larger volume. The books are markedly different in the level at which they are pitched: Rethlingshafer's is distinctly elementary or intermediate, at best, whereas the Cofer and Appley volume could easily be used as a graduate level text; in their larger book, Cofer and Appley can proceed in a more leisurely manner and, as a result, they have been able to provide a number of valuable summaries of earlier ideas and research as a means of introducing many topics—this is best exemplified in the chapter entitled "Motivation in Historical Perspective." Even when, as is often the case, they are advocating a particular formulation, Cofer and Appley have a solidity and judiciousness that Rethlingshafer does not have; to this reviewer, her presentation too often has either a strident, argumentative quality or else is far too bland considering the complexity of many of the issues—a difference which may in part be attributed to the smaller scale of Rethlingshafer's effort.

The two books are in no sense of the word competitors. Though they both cover similar territory and offer accounts of current thinking about motivation, the Cofer and Appley book is by far the more scholarly. This latter quality is enhanced by the contrast between their sober but quite skillful approach to their task and the frequency with which Rethlingshafer's text jarred this reviewer by commonplace observations, illustrations, and brief homilies presumably designed to illuminate or humanize otherwise difficult and technical material. Both books have fine bibliographies and indices, in each case amounting to well over 10% of the total pages; the reviewer approves the combined bibliography-author index that Cofer and Foley use. Rethlingshafer has used "insets" (which are dark type-face sections) to good advantage as a device to present extensive quotations or for summarizing a particular experiment or piece of logic. Finally, Rethlingshafer's book is more selective and sketchy so that a student is less likely to realize how many byways and interesting alternatives he has been led around—but that is the difference between an elementary and an advanced book!

University of Oregon
RICHARD A. LITTMAN

Behavioral Individuality in Early Childhood. By ALEXANDER THOMAS, HERBERT G. BIRCH, STELLA CHESS, MARGARET E. HERTZIG and SAM KORN. New York University Press, 1963. Pp. xii-135. $4.75.

This is a disappointing first report of an unusually interesting research project. The project is unusual for many reasons. First of all, it is a longitudinal study which knows where it is going. It is explicit about its goals, its methodology, and the behavior being studied. This is a refreshing change from those studies which amass data at different levels of behavior in a catch-as-catch-can manner, with the hope that answers to important questions will somehow stand naked and revealed.

The second methodological asset is the evidence that mothers can accurately report their children's behavior if (1) the behavior is recent and (2) if it is objective and observable. The group is Watsonian in its insistence that the basic datum of psychology are the observable facts of human behavior. Their finding that mothers can validly portray such facts puts researchers in their debt.

The content of the project is more exciting than the methodology. Thomas, et al. are interested in general characteristics of behavior which reveal themselves in the first six months of life and which remain relatively constant. Some have called such characteristics constitutional, or temperament, or the expressive aspects of behavior. Whatever the label, the focus is on the "How" of behavior, not the "What"—how the infant responds to a wide range of experience, not his particular reaction to a discrete experience such as feeding, toileting, etc. The authors rightly point to the neglect of this important area during the past few decades. They imply that the lacklustre results of attempts to relate maternal attitudes or child care practices to personality development, may be due to a failure to evaluate the child's individualized patterns of responding.

The other exciting feature of the project is that in choosing the facts of infant behavior as the principle data, the investigators have turned their backs upon the two most popular variables in developmental research intrapsychic variables, includes affective states and fantasies, and interpersonal variables. Because of the clarity of their intent, these investigators will be able to show just how far one can go if one starts with the overt behavior of the infant. From this we can learn when intrapsychic and interpersonal factors are excess baggage and when they need to be called into play. This will be a giant step forward.

It is just because the project has so much to recommend it, that this report is a disappointment. Much of the material has appeared in journals and the authors have used the increased elbow room of a monograph unwisely.

Let us start with some conceptual and technical criticisms. Thomas et al. describes nine categories of behavior—activity level, rhythmicity, approach or withdrawal, adaptability, intensity of reaction, threshold of responsiveness, quality of mood, distractibility, attention span and persistence—which are observable in the first six months of life. The heart of their study is their claim that these reactions remain consistent over a two year period, the span of the project covered by this monograph. They imply that such consistency is strong evidence for built-in individuality. However, they do not consider the possibility that the consistency may be due, in part, to the fact that only middle class families were studied. They also ignore the possibility that the mother may play a significant role in perpetuating certain kinds of behavior; e.g., that the infant's mood or adaptability may have been determined by the manner in which he was handled.

Ironically, the investigators must be aware of the first criticism because they are in the process of studying infants from a low socio-economic class. They also have data on maternal handling and can therefore answer the second criticism. By not raising the issue of alternative explanations of their data, the authors sell themselves short.

Their technical evidence for consistency of reaction patterns over time is equally open to question. The table of correlations for their third and best measure of consistency shows that most of the significant predictions are from one six month period to the next; when long range predictions are made from six months of age to two years of age, only two of the categories of reaction patterns reach significance. Such results are reminiscent of those from infant intelligence tests, which can predict what lies in the immediate future, but which are useless for longer spans of time. Thus, both at the level of design and data analysis, the group has not presented a totally convincing case for their main thesis concerning the stability of initial patterns of reaction.

Next, the authors devote a disproportionate amount of space to pointing the finger of shame at other longitudinal studies and at psychoanalytically oriented research. Their criticisms are well taken, but their energy is misused. It would have been better for them to present a more searching and thoroughgoing account of the theoretical and practical issues involved in their own effort to understand the "How" of behavior.

Finally, the chapter on practical implications does not belong in the monograph, either in spirit or in content. It is full of pat illustrations of how, by taking a child's individual reaction patterns into account, problems in sleeping, toileting, etc., can be managed successfully. Pronouncements are ex cathedra and innocent of evidence, and the freewheeling style is out of keeping with the thoughtful and hardheaded approach of the rest of the monograph.

In sum, the research of Thomas et al. is too important to be reported so unsatisfactorily. It is hoped that future publications will do full justice to the scope of the team's ideas and data. It is also hoped that having gotten their criticisms of other studies out of their systems, the group will delve more deeply into the theoretical implications of their own work. Finally, it is hoped that, if a chapter on practical considerations is included at all, it contains the same careful attention to evidence and design that characterizes the heart of the project.

University of Pennsylvania — Charles Wenar

Some Views on Soviet Psychology. Edited by R. A. Bauer. Washington, D.C., American Psychological Association, 1962. Pp. ix, 285.

This is a collection of papers written by American psychologists about their visits to the Soviet Union during the summer of 1960 as part of a project sponsored by the American Psychological Association and supported by the Human Ecology Fund. The eight chapters begin with an historical introduction by Alexander Mintz. The other topics included are: problem solving and thinking by Walter Reitman, personality development and socialization by Urie Bronfenbrenner with comments by Otto Klineberg, Soviet mental health by Henry and Jema David, developmental psychology by Yvonne Brackbill, industrial and educational psychology by Edwin Fleischman, general psychology and psychophysiol-

ogy by Neal Miller, Carl Pfaffmann, and Harold Schlosberg, and Soviet life and psychology by Gardner and Lois Murphy.

The scope of the papers differs considerably depending on the length of the individual visits and the particular aspects of Soviet psychology which were seen. It was intended that the different specialties of the visiting psychologists would permit a reasonably precise appraisal of the status of Soviet psychology. Mintz was the one contributor who was not reporting on a visit. His knowledge of Soviet psychology is, however, well attested to by his articles about it in the *Annual Review of Psychology*. His introduction does give the reader a good perspective with which to consider the other chapters. He documents an amount of political interference in Soviet psychology which is disappointing but he is generally quite optimistic about the possibilities of coöperation between western and Soviet psychologists.

Reitman describes three different approaches to the study of thinking taken by the Soviets: experimental, pedagogical, and cybernetic. He also notes that many studies purporting to involve examination of "higher nervous activity" are behavioral in nature and only relate to physiology by assumption. But he also does a service for the reader by abstracting basic physiological assumptions inherent in the Soviet interpretation of experiments on higher nervous activity. Bronfenbrenner points out the emphasis on the collective as a socializing agent in the Soviet Union, describing in some detail the work of Makarenko, an early Soviet educator, who put such an emphasis into practice. The Davids describe the administrative organization of Soviet mental institutions as well as type of care given. They comment on the generally organic approach to mental illness taken by Soviet psychologists. They as well as Bronfenbrenner indicate that the Soviets recognize Freud's unconscious but have downgraded it as being relatively unimportant. They also note in common that social psychology is the prerogative of the Party.

Brackbill's contribution is probably the most detailed description of a particular area of investigation, being based on two trips to the Soviet Union as well as a familiarity with the technical literature. Her chapter includes topics ranging from research with infants, through cognitive development of preschoolers to clinical psychology with children and has an appendix with a detailed description of apparatus. Fleischman comments on industrial and educational psychology in Moscow, Leningrad, and Tbilisi. At that time the main effort in industrial psychology was in Moscow where work was concentrated on labor activity (work methods, structoring of skills, monotony, etc.), occupational training in adults and in children. On the basis of visits to classrooms and pioneer centers as well as laboratories Fleischman was quite impressed with the great commitment to education on the part of the whole Soviet society.

Miller, Pfaffmann, and Schlosberg make precise comments about the training of psychologists and many of the research facilities in Moscow, Leningrad, Sukhumi (primate colony), and Tbilisi. It is impossible to single out any particular points to highlight. Their report covers the areas of work in the Soviet Union corresponding most closely to what we would call comparative, experimental, and physiological psychology. They do make some summary comments of interest: (1) Soviet interest in the reticular formation is widespread; (2) *EEG* is often used simultaneously with conditioning and verbal reports; (3)

conditioning is a generic term for all kinds of learning; and (4) there is an interest in individual differences characterized by a Pavlovian typology. (Their report also includes an appendix on a visit to Warsaw.)

The Murphys describe much more of the Soviet way of life in their article, noting for example, how children are treated and the kinds of facilities there are for children. This is not a simple travelogue, however, for they integrate such observations nicely into accounts of the theoretical aspects of Soviet psychology. They also include detailed descriptions of work done with the handicapped and emotionally disturbed children.

In general the papers are very full and quite accurate in view of the fact that in most cases it was necessary to operate through an interpreter. Of course, as most of the authors stress, their impressions just scratch the surface of what is going on. Nevertheless, the book still provides a service (although it's almost five years since the visits were made) for those who want a nodding acquaintance with Soviet psychology. Since the book was written there has been a tremendous increase in engineering psychology in the Soviet Union with an accompanying interest in mathematical techniques (but still little statistics) and a smaller but probably significant increase of interest in social psychology. It is to be hoped that soon Soviet work will be well enough known in the United States to be included as part of the normal topical literature reviews rather than requiring such an international tourist approach.

A final note of interest is that A. R. Luria received this book from the Soviet point of view (*Voprosy̆ psikhologii,* 1963, 9(3), 147-148). He concludes that the book is generally quite an accurate account and is written objectively and with good will.

University of Minnesota HERBERT L. PICK, JR.

Multivariate Procedures for the Behavioral Sciences. By WILLIAM W. COOLEY and PAUL R. LOHNES. New York, John Wiley and Sons, 1962. Pp. x, 211. $6.75.

Cooley and Lohnes present some of the more useful multivariate procedures (as well as some computationally demanding univariate ones) appropriate to statistical analysis of data arising frequently in the behavioral sciences. The special value of multivariate procedures lies in their potential to test hypotheses at an appropriately general level by the simultaneous consideration of sets of variables and thereby providing unique information not found by less comprehensive methods. For example, there may be a substantial overall relationship revealed by canonical correlation between two different *sets* of variables, each representing a particular domain such as early home environment and present orientation towards people. Separate comparisons of each variable in one set with each variable in the other set would give no indication of the extent of over-all relationship. Or canonical correlation might show that there is no possible combination of a number of satisfaction measures which is significantly related to any combination of several performance measures, testing a general hypothesis about the over-all relation between the domains of satisfaction and performance as spanned by the available measures. In this case, an insignificant canonical correlation emphasizes the inappropriateness of searching for relationships between pairs of individual variables.

This book begins with a brief summary of multivariate procedures and the questions to be answered by each method. The procedures include: multiple and can-

onical correlation; multivariate analysis of variance and covariance; univariate analysis of variance and covariance for 2- and 3-way designs; multiple-discriminant analysis and classification procedures; factor analysis, rotation, and computing of composite scores. In several instances useful significance tests are referred to and included in the programs.

Each procedure is presented very concisely and the price for this brevity sometimes is an uneven level of exposition. There is some reference to the mathematics involved, and sometimes geometric interpretations are used in an attempt to clarify explanations. The descriptions of canonical correlation, multivariate analysis of variance, and discriminant and classification procedures are particularly good. In most cases, computed examples contribute to a better appreciation of the special advantages gained with multivariate analysis. The selected references suggested for more detailed information on particular techniques are excellent and include expository as well as rigorous treatments.

Flow diagrams are presented as a guide for programming each technique for any computer and these all appear to be correct. FORTRAN-coded programs are provided for IBM 704-709 computers and apparently there are no major errors in the listings, although for obvious reasons these were not checked in detail. It is unfortunate that no estimates of computing time were given. It should be assumed that anyone using programs punched from these listings should run a test-case using data for which results are known.

The unique feature of the book is its attempt to facilitate the implementation of multivariate analyses. The tested programs are ready to run, with perhaps minor modifications required for particular installations. Usually the programs will handle up to 50 variables, and the analysis of variance routine for 2- and 3-way designs is especially versatile. But computing routines become outdated even faster than computers. Persons interested in using any of these techniques for large-scale studies where efficiency is important, and those interested in having available all the techniques discussed in the book are well advised to obtain a copy of an updated program tape (for details contact Kenneth Jones at the Harvard Computing Center). Obtaining programs from tape also has the advantage of avoiding errors that often arise when keypunching many statements from a listing (*e.g.* I and 1 are easily confused). A revision of the book is planned to include updated subroutines.

Several minor errors were found in the text. On p. 34 *df* for the denominator of the *F*-test for R^2, and the corresponding term in the numerator on the right side of the equation, should be N-m/(not N — m —2). (The flow diagram and program do not contain this error.) On pp. 44-45 *df* for χ^2 for the canonical correlation should be *36*, (not *24*). On p. 118 the percentage of total discriminating power of the *j*th discriminant function, (not the *i*th), is represented by the second equation. On p. 137 the text following Decision Rule I should refer to Group 1 χ^2, (not χ_2). Throughout the example beginning on p. 165 the third variable should be $\overrightarrow{W}$ (not $\underline{W}$). On p. 187 it does not seem reasonable that the approximate time for computing roots and vectors with HDIAG is n^3 sec. Possibly this should appear as n^3 *milli*seconds.

It was necessary for the authors to select among alternative approaches to some of the techniques, such as factor analysis and rotation, and the selections were influenced by their extensive experience in data-analysis. They recognize that their

recommendations may not represent the most elegant way of conducting factor-analysis, for example, but they feel that most of the suggestions for refinements tend not to affect the psychological interpretation of the factor-solution. This is a reasonable position but it is worth repeated emphasis that none of these techniques can be used properly by the uninformed. This book certainly will assist in generating numbers summarizing data but it does not remove the user's responsibility to understand the assumptions underlying a given model and their influence on the proper interpretation of results. This warning is sounded once in the book but perhaps it bears repeating in an attempt to discourage uncritical applications.

Tne most lasting value of this book may be in its use in introducing students to the implementation of multivariate analyses. It will not serve as a self-contained text but the material provided serves as an excellent link between theory and application. It will contribute substantially to making available to a wide audience much of what is possible in the area of multivariate analysis.

Educational Testing Service L. M. KENDALL

Color: A Guide to Basic Facts and Concepts. By R. W. BURNHAM, R. M. HANSE, and C. JAMES BARTLESON. New York, John Wylie & Son, 1963. Pp. xii, 250.

This book presents the final report of the Committee on Basic Elements of Color Education, Problem 20, of the Inter-Society Color Council. It is a compilation of a large number of 'facts,' frequently in the form of a single sentence without context. The statements are grouped under a combined numerical-alphabetical system that sometimes runs to as many as 6 places. No exposition of the system is given. Apparently, it is expected to be obvious, but the reader is left to wonder whether, for example, Item 3.4.3.a.l.a is a lower order of 'fact' than Item 3.4.1.

The Introduction presents the views of the writers concerning their subject-matter. Those constitute a somewhat naïve epistemology which is expressed by the statement: "Color is, first of all, a word that makes it possible for us to tell others something about what we are aware of when we look at the things about us." A word, obviously, though definable, can not be restricted in its usage. 'Color' as a word is no exception and therefore means many things to many people. Similar introductions have been used before, and it appears to the reviewer that the present exposition achieves less clarity than Hering's statement of the problem in 1878. Bridgeman has given us a useful definition of fact in terms of the operations followed to produce specific data. We would do well to rely on that concept.

Undoubtedly most of the items or 'facts' are substantially correct, especially when they consist of data in the form of tables and graphs. It is beyond the scope of this review to consider the validity of separate statements. Such validation may need to be undertaken by the individual user. This will not be an easy task because very few specific references are given, and many statements are generalizations or interpretations which reflect the epistemological assumptions of the writers. The result leads sometimes to wordy confusion, sometimes to downright error. For example, the interpretation of Fig. 3.18 in Item 3.5.2. as showing the "color zones of the retina" is in error. The figure is, in fact, a map of the visual field, obtained with a campimeter using a flat projection field, not with a perimeter having a spherical field. The scale-units applied to the concentric circles are quite arbitrary and do not represent equal distances or angles projected on the retina. The error in this case

may have been the unfortunate selection of the source. The referenced publication copied the diagram without giving credit to the originator. The source publication (Boring, Langfeld and Weld, *Foundations of Psychology,* 1948) gives a more adequate account of the operations involved in obtaining this map of the visual field.

In spite of the wordy involvement with color as awareness and the absence of a clearcut statement of operations required for obtaining color-data such as color qualities and color discriminations, the authors have compiled a large assortment of figures and formulae referring to many aspects of color stimuli. There are, also, several interesting plates that demonstrate some color phenomena, mostly involving spatial effects. An increase in this section might, indeed, have justified the title of the book.

New Brunswick, New Jersey FORREST LEE DIMMICK

Spatial Development. By P. VEREECKEN. Groningen, J. B. WOLTERS, 1961. Pp. vii, 152. f17.50.

Spatial Development is an unusual little book on the understanding of tasks involving the perception and performance of line drawings and block-building activities by children ranging in age up to seven years. Vereecken takes as his goal the explanation of "constructive-praxic" activities, *i.e.* "voluntary acts." Voluntary as used throughout the book means only that the child is aware of what he is going to do before doing it.

Vereecken tested many tasks and the ages at which they could be performed. The tasks as well as responses of his children are well illustrated, so that his facts are given neatly to the reader.

His explanations of development and of other factors required for the child's understanding of the performance of the tasks are complex. His account relies heavily on a motor theory of learning which incorporates the ego, the Gestalt laws of organization (unexplained) as well as a heterogeneous assortment of dynamic determinants such as "closing in forces" and "fears of empty spaces" on test-cards.

Most of his tasks required children to replicate line drawings or block patterns. One of the difficulties of research in this area is determining when a pattern has been successfully replicated so that an experimenter can conclude that a child really understood what he was doing. Differential rates of motor development also complicate the situation, as noted by the author.

This book continues the tradition of small-sample research as the basis for theory. This reviewer asserts that the social significance of theory requires large samples for testing its generality and that investigators who propose to be theorists should run their preliminary samples and then check their important findings on large samples. This the author did not do. Consequently, his statements of ages at which children can perform and understand tasks require much more investigation before they can have any status other than assumed populational generality.

Researchers who are interested in problems of intelligence or in processes of problem solving required for the replication of line drawings or block patterns should read this book so as to become familiar with his data. Those who intend to work in the area will find the bibliography of assistance even though references are occasionally made to authors not listed in the bibliography.

Cornell University OLIN W. SMITH

The Design of Electric Circuits in the Behavioral Sciences. By TOM N. CORNSWEET. New York, John Wiley and Sons, 1963. Pp. xi + 329. $8.95.

This book is addressed to the psychologist who needs some electrical skills in the conduct of his research. It acquaints him with basic circuit elements—switches, relays, capacitors, transformers, and so forth—and demonstrates their use in the presentation and timing of stimuli, sensing the movements or positions of animal subjects, and many other experimental applications. The book does not pretend to give definitive circuitry for any particular application; it attempts rather to establish enough comprehension of various simple circuits to permit the experimenter to design his own—or, if his problem is especially complex, to discuss design problems critically with a technician.

In my experience, it is fairly easy to teach novice research assistants to use schematics, and it is also fairly easy to show them how particular electrical devices work. To bridge the gap from a schematic to a completed working circuit, however, seems extraordinarily difficult. The great virtue of this book is that it makes the transfer from the abstractions of a schematic to the operation of the equipment as simple as possible. Cornsweet's technique is to intersperse photographs of actual components and sketches of their applications with circuit diagrams, electrical theory, and practical wiring instructions. The book presents a series of problems that are best solved by actually wiring up a set of recommended components and seeing that the thing works—which comes as an exciting revelation to the novice.

Several chapters of the books are quite specific. For instance, there is a good section on soldering, and an excellent chapter is devoted to the problem of how to shock a rat. Other chapters give an overview of large classes of circuits and equipment, such as amplifiers and oscilloscopes. On the whole, the treatment is clear and easy to follow. My only criticism is that the topic of complex switching circuits has a moderately difficult introduction and moves too fast for the complete novice.

The book ought to be especially useful to graduate students with a limited technical background. If they actually work through the problems, they will learn a great deal of general value, and extinguish their fears of electrical apparatus. I have used it with undergraduate students with good results.

Swarthmore College J. A. NEVIN

Biochemistry and Behavior. By SAMUEL EIDUSON, EDWARD GELLER, ARTHUR YUWILER and BERNICE T. EIDUSON. Princeton, New Jersey, D. Van Nostrand, 1964. Pp. xii, 554. $15.00.

This is an important, useful and difficult book. The authors are three biochemists and a psychologist. Their aims are biochemical and behaviorial. On the biochemical side they have succeeded admirably. From about 2000 references they have managed to summarize some of the more recent data on how chemicals commonly discussed in connection with behavior are built up, broken down and localized within the brain. This is probably the best single source for this information. They guide the reader skillfully through the biochemistry: main themes appear, and qualifications and exceptions appear as such. This is serious biochemistry and not the place to learn the biochemistry which has never been learned. But this

book can be recommended highly to anyone who wants to find out the latest word on these chemicals and get important leads in the research literature.

When it comes to behavior they are less successful. For one thing the behavior studies are a crazy-quilt of poor studies, uncontrolled observations, and unreliable data. The authors point this out better than anyone else could and they offer some caveats for future work. The treatment of the behavioral data is, however, still disjointed. They seem not to have tried as they did in the biochemistry, to find the unifying thread or the meaning latent in several partial studies. They end up cataloguing rather than ordering the field. Still, they perform the enormously useful function of surveying this vast literature in chapters on energy metabolism, neurohumors, hormones, chemical factors in mental illness, psychologically active drugs and biochemical genetics. The authors have presented these studies broadly and with care. The major findings and theories are all presented.

In addition to its stimulus to present research, the book ought to have a second long-range function of requiring interdisciplinary workers to recognize that we are now at a point where our behavioral observations ought to require more effort and planning than biochemical ones. Elegant biochemical determinations on crude behavioral are going to leave us with just another inconclusive study. Gresham's Law is at work here: bad behavior will drive out good biochemistry. If this book serves to advance this lesson, it will have served admirably a present and a future purpose.

Yeshiva University ALLAN C. GOLDSTEIN

On Fundamental Measurement in Psychology. By GUNNAR GOUDE. Uppsala, Almqvist and Wiksell, 1962. Pp. 176. Sw. kr. 18.

After philosophizing for several chapters on the nature of measurement, Goude appears to conclude that, since physics has used a physical operation of addition as the basis for fundamental scales, psychology ought to use a corresponding psychological operation of addition. Additive operations are proposed for psychophysical continua and for ability continua. For psychophysics, psychological addition is held to occur when the subject is presented with two standards (a and b), and then told to adjust a variable c until it appears as great as the sum of a and b. There is nothing really wrong with this procedure except its interpretation as a logical parallel of physical addition. For psychological addition of ability, Goude suggests simply having subjects with abilities x and y work on single-factor test problems cooperatively—their cooperative score then being interpreted as a function of the sum of the two abilities.

The major share of the monograph is devoted to empirical "verification" of these procedures. For psychophysics, this involves presentation of a host of the author's own small experiments using the addition method, along with many of the standard judgment methods, on the attributes of weight and angle. Although he concludes that the methods used all yield the same fundamental scale, the results are not at all convincing. The psychophysical experiments not only are too limited in scope but also involve highly suspect *ad hoc* modifications to the methods used. For ability, results of a preliminary study involving number series and opposites are reported. Although Goude again concludes that the data do not refute the ideas of the addition method, there are just too many things wrong with the notion of cooperation as addition of ability to allow either the theoretical

argument or the empirical results to be taken seriously. Short chapters on measurement in plant and animal psychology involving the scaling of luminosity for oats and mice are also included.

Lincoln Laboratory WARREN S. TORGERSON
Massachusetts Institute of Technology

Motivation: A Biological Perspective. By JOHN L. FULLER. New York, Random House, 1962. Pp. 116. $1.45.

Any attempt to provide a biological perspective of motivation at this point must necessarily be incomplete and, at points, superficial. Fuller assumes that motivation is a necessary intervening variable the definitions of which are the production of responses (p. 4), "stabilized organization of directed behavior," (p. 83) "the maintenance of a disposition to respond in a particular direction until reinforcement occurs," (p. 88) or "the psychological consequence of a balanced state of the nervous system" (p. 91). The diversity in definition of the concept probably reflects the state of the field rather accurately.

Initially, motivation is explained in terms of drive-reduction and survival theory. Biological needs, integrative needs, and innate and acquired drives serve as major categories for this discussion. Interestingly enough, a limited number of acquired drives, all of which are *emotions,* is postulated. When discussing the physiology of motivation, particularly the role of the nervous system, the author shifts the emphasis to activation theory.

The discussion of social motivation, conflict, and human motivation, not surprisingly, is rather weak since the biological foundations of behavior in these categories have not been investigated with the same intensity as the physiology of hunger, thirst, and sex. Fuller is willing to conjecture, however, that the principles of human motivation will be found to fit the paradigm of need, primary drive, acquired drives, expressive drives, etc.

The brief introductory statement of the problems of the measurement of motivation is excellent, assuming that this book is intended for use in a first course in psychology. The misuse of time, rate, and intensity measures is discussed and measures of "resistance to change" and "persistence toward goals" are recommended as more adequate.

Bucknell University WENDELL SMITH

Causalité, Permanence et Réalité Phénoménales: Études de Psychologie Expérimentale. By A. MICHOTTE and collaborators. Louvain, Publications Universitaires, 1962. Pp. 612. FB 500.

This is an important collection of twenty-six papers by Professor Michotte and nine collaborators: M. Yela, L. Knops, A. Coen-Gelders, W. J. M. Levelt, G. Crabbe, A. C. Sampaio, L. Burke, A. J. Glynn, and M. R. Phemister. A few of the papers are new but most are reprinted from diverse European sources, their original dates of publication ranging from 1941 to 1960. The collection will be especially welcome to American readers because it gives a sense of the wide sweep of Michotte's work in two ways: first, his achievement of a *differentiated* view of the experience of causality, not as a single kind of perceptual impression but a group of related ones, varying among themselves in an orderly way depending on stimulus conditions; secondly, his examination of the causal problem in relation to a variety of other

issues, such as the phenomenal permanence and the phenomenal reality of objects, and the interrelations among perception, cognition, and language. Just as a sampler, the papers by Knops and Sampaio on phenomenal permanence are very interesting for their analyses of the stimulus conditions which control the impression that vanished objects maintain their existence; Professor Michotte's paper on the influence of past experience on perception represents a valuable clarification of his near-nativist views on that subject.

New School for Social Research HOWARD E. GRUBER

Personality: An Experimental Approach. By R. W. LUNDIN. New York, The Macmillan Company, 1961. Pp. xi, 450. $5.75.

This book could more precisely be titled *Personality: A Learning Theory Approach,* or, even, *A Skinnerian Approach.* In the first two chapters the author describes alternate theories of personality, both historical and current, and states his own position, which is that personality is the "unique behavior equipment which one has acquired through a past history of learning." He then spends several chapters on explaining and illustrating with laboratory experiments such concepts as classical and operant conditioning, extinction, reinforcement schedules, discrimination, deprivation and satiation, etc.; and relating these to simple human behavior. More complex human responses such as escape, anxiety defense and conflict, as well as neurotic and psychotic behavior, are also explained within the learning theory framework with many illustrative experiments using both animal and human subjects. Usually the relation between the author's approach and a psychoanalytic one is made explicit. In a final chapter he describes some possible applications of his approach to psychotherapy.

The book is well organized and simply and clearly enough written to be entirely suitable for the less advanced student. The instructor might want to supplement it with material presenting other points of view.

Hunter College ELIZABETH G. FRENCH

Industrial Psychology By LAURENCE SIEGEL. Homewood, Illinois, Richard D. Irwin, 1962, Pp. xvi, 414. $6.95.

This book is organized around three subgroups of people, workers, management, and consumers. This about covers the waterfront and that is what this text attempts to do. Unfortunately this procedure leads to rather superficial coverage of a multitude of topics ranging from scientific method to selling and advertising. Perhaps because of this skimming approach, statements are made with little or no evaluative evidence. This is especially true in the sections on selling and advertising. In other instances uncritical statements are made at one point in the book and it is not until later that questions are raised concerning these remarks. An example of this is the author's comment concerning motivation research, where on page 348 he describes how useful this approach can be and it is not until page 360 that he raises the question of validity. It should be pointed out, however, that Siegel is concerned with the criterion problem and develops this well throughout the text.

The book seems, to this reviewer, uneven in difficulty level. Perhaps the problem here is the lack of clarity concerning the audience to whom it is addressed.

New York University JOSEPH WEITZ

Learning. By SARNOFF A. MEDNICK. Englewood Cliffs, New Jersey, Prentice-Hall, 1964. Pp. vii, 118. $3.95 (Paper, $1.50).

This book is a competent survey of animal and human learning designed for the beginning student. It covers terminology and concepts, conditioning, complex habits, motivation, transfer, and memory. The book is written in a popular style—perhaps too popular—and is not hard to understand; the general approach to most of the topics is to start with an example from everyday life and lead into a discussion of controlled experimental studies. While some classic studies are cited, the author also tries to indicate to the student some of the excitement and enthusiasm engendered by recent developments in the field. While this book does not provide intensive coverage of any one field, it does provide an overview (from an S-R vantage point) of the general learning area plus brief accounts of a number of representative and important experiments.

University of Missouri BENNET B. MURDOCK, JR.

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review.)

ABRAHAMSEN, DAVID, *The Psychology of Crime.* New York, Wiley, 1964. (First published 1960). Pp. xii, 358. $1.95 (paper).

ANDERSON, RICHARD C. and AUSUBEL, DAVID P. (eds.), *Readings in the Psychology of Cognition.* New York, Holt, Rinehart and Winston, Inc., 1965. Pp. v, 690. $7.50.

ANGYAL, ANDRAS; HANFMANN, E.; and JONES, R. M. (eds.), *Neurosis & Treatment· A Holistic Theory.* New York, John Wiley & Sons, Inc., 1965. Pp. v, 328. $8.50.

BACHELARD, GASTON (trans. by ALAN C. M. ROSS), *The Psychoanalysis of Fire.* Boston, Mass., Beacon Press, 1964. Pp. viii, 115. $3.95.

BANTON, MICHAEL, *The Policeman in the Community.* New York, Basic Books, 1965. Pp. xiv, 276. $5.50.

BARDILL, DONALD R. and RYAN, FRANCIS J., *Family Group Casework. A Casework Approach to Family Therapy.* Washington, D.C., Catholic University of America Press, 1964. Pp. vi, 58. $1.50 (paper).

BARNES, C. D. and ELTHERINGTON, L. G., *Drug Dosage in Laboratory Animals: A Handbook.* Berkeley, Univ. of California Press, 1964. Pp. 302. $8.00 (paper).

BARTON, HELEN B., *Nervous Tension, Behavior and Body Function.* New York, Philosophical Library, 1965. Pp. 336. $6.00.

BERGLER, REINHOLD, *Psychologische Marktanalyse.* Bern und Stuttgart, Verlag Hans Huber, 1965. Pp. 352.

BERLYNE, D. E., *Structure and Direction in Thinking.* New York, John Wiley & Sons, Inc., 1965. Pp. v, 378. $8.95.

BLUMENTHAL, MURRAY (ed.), *The Denver Symposium on Mass Communications Research for Safety. A Critical Review of the Literature and a Proposed Theory by H. A. Mendelsohn.* Chicago, National Safety Council, 1954. Pp. viii, 288. $5.00 (paper).

BUGENTAL, J. F. T., *The Search for Authenticity: An Existential-Analytic Approach to Psychotherapy.* New York, Holt, Rinehart, and Winston, Inc., 1965. Pp. vii, 437. $7.95.

CAPPON, DANIEL, *Toward an Understanding of Homosexuality.* Englewood Cliffs, N.J., Prentice Hall, 1965. Pp. xi, 302. $6.95.

CARTWRIGHT, ANN, *Human Relations and Hospital Care.* New York, The Humanities Press, Inc., 1964. Pp. v, 262. $6.00.

CASTELL, ALBUREY, *The Self in Philosophy*. New York, The Macmillan Company, 1965. Pp. iii, 122. $1.95 (paper).

CASTELNUOVO-TEDESCO, PIETRO, *The Twenty-Minute Hour: A Guide to Brief Psychotherapy for the Physician*. Boston, Little Brown and Company, 1965, Pp. vii, 184. $5.95.

COSTIGAN, GIOVANNI, *Sigmund Freud: A Short Biography*. New York, The Macmillan Company, 1965. Pp. xiv, 306. $4.95.

DENES, GYULA, *Change your Life with Positive Action*. New York, Frederic Fell, Inc., 1964. Pp. 211. $4.95.

DRILLIEN, CECIL M., *The Growth and Development of the Prematurely Born Infant*. Baltimore, Williams & Wilkins Company, 1964. Pp. xv, 376. $9.50.

DUNN, L. C., *Hereditary and Evolution in Human Populations* (Rev. ed.). New York, Atheneum, 1965. (First published 1959.) Pp. viii, 155. $1.95 (paper).

EBBINGHAUS, HERMANN, *Memory. A Contribution to Experimental Psychology*. New York, Dover Publ., 1964. (Reprint of 1913 English translation.) Pp. xvi, 123. $1.50 (paper).

EDELMAN, MURRAY, *The Symbolic Uses of Politics*. Urbana, Illinois, Univ. of Illinois Press, 1964. Pp. 201. $5.00.

FARBEROW, NORMAN L. and SHNEIDMAN, EDWIN S. (eds.), *The Cry for Help*. New York, McGraw-Hill Book Company, 1961. Pp. ix, 398. $3.45 (paper).

FARNSWORTH, P. R. (ed.), *Annual Review of Psychology, Vol. 15, 1964*. Palo Alto, California, Annual Reviews, Inc., 1964. Pp. vii, 629. $8.50.

FISH, FRANK, *An Outline of Psychiatry. For Students and Practitioners*. Baltimore, Md., Williams & Wilkins, 1964. Pp. 270. $6.95.

FISH, FRANK, *Clinical Psychiatry for the Layman*. Bristol, England, John Wright & Sons, 1963. Pp. vii, 59. $2.95 (paper).

FLORY, CHARLES D. (ed.), *Managers for Tomorrow*. New York, The New American Library, 1965. Pp. 352. $5.95.

FREEMAN, RUTH ST. JOHN; and FREEMAN, HARROP A., *Counseling: A Bibliography (with Annotations)*. New York, Scarecrow Press, 1964. Pp. xxxv, 986. $22.50.

FRIEDLANDER, ERNST, *Psychology in Scientific Thinking*. New York, Philosophical Library, Inc., 1965. Pp. 143. $4.75.

FULTON, ROBERT (ed.), *Death and Identity*. New York, John Wiley & Sons, Inc., 1965. Pp. v, 415. $6.95.

GARRISON, KARL C., *Psychology of Adolescence*. (6th ed.) Englewood Cliffs, N.J., Prentice-Hall, Inc., 1965. Pp. xii, 487. $7.95.

GOWERS, W. R., *Epilepsy and Other Chronic Diseases: Their Causes, Symptoms, and Treatment*. New York, Dover Publ., 1964. Pp. xi, 255. $1.85 (paper).

GROSSMAN, CARL M. and GROSSMAN, SYLVIA, *The Wild Analyst: The Life and Work of Georg Groddeck*. New York, George Braziller, Inc., 1965. Pp. 222. $5.00.

GROUP FOR THE ADVANCEMENT OF PSYCHIATRY, *Urban America and the Planning of Mental Health Services: Symposium No. 10*. New York, Group for the Advancement of Psychiatry, 1964. Pp. 387-518; 144 pp. $1.50 (paper).

HAAS, KURT, *Understanding Ourselves and Others*. Englewood Cliffs, New Jersey, Prentice-Hall, Inc., 1965. Pp. xi, 370. $9.25.

HOLLAND, GLEN A., *Fundamentals of Psychotherapy*. New York, Holt, Rinehart and Winston, Inc., 1965. Pp. vii, 308. $6.95.

HULL, CLARK L., *A Behavior System: An Introduction to Behavior Theory Concerning the Individual Organism*. New York, Wiley, 1965. (First published 1952). Pp. ix, 372. $2.25 (paper).

HULTZEN, LEE S.; ALLEN, JOSEPH H. D., JR.; and MIRON, MURRAY, S., *Tables of Transitional Frequencies of English Phonemes*. Urbana, Ill., Univ. of Illinois Press, 1964. Pp. 49 plus unpaged tables. $5.00.

ILG, FRANCES L. and AMES, LOUISE BATES. *School Readiness: Behavior Tests Used at the Gesell Institute*. New York, Harper & Row, 1965. Pp. xiii, 396. $7.95.

KETCHUM, J. DAVIDSON, *Rubleben: A Prison Camp Society*. Toronto, Canada, Univ. of Toronto Press, 1965. Pp. vii, 397. $7.50.

KORNHAUSER, ARTHUR, *Mental Health of the Industrial Worker.* New York, Wiley, 1965. Pp. vii, 354. $7.95.

KRIS, ERNST, *Psychoanalytic Explorations in Art.* New York, Schocken Books, 1964. Pp. 358. $2.95 (paper).

LANGHOFF, PETER (ed.), *Models, Measurement and Marketing.* Englewood Cliffs, N.J., Prentice-Hall, Inc., 1965. Pp. 216. $6.95.

LAUZUN, GERARD, *Sigmund Freud: The Man and His Theories.* New York, Paul S. Eriksson, Inc., 1965. Pp. 224, $5.00.

LEVINE, DAVID (ed.), *Nebraska Symposium on Motivation, 1964.* Lincoln, Nebraska, Univ. of Nebraska Press, 1964. Pp. x, 284. Cloth, $5.50; paper, $2.50.

LUCE, R. DUNCAN; BUSH, ROBERT R.; and GALANTER, EUGENE (eds.), *Readings in Mathematical Psychology. II.* New York, John Wiley & Sons, Inc., 1965. Pp. v, 568. $8.95.

MACCOLL, SYLVIA HAZELTON, *Structure and Development of Phenomenal Reality: A Psychological Analysis.* New York, Exposition Press, 1965. Pp. 222. $5.00.

MAHER, BRENDAN A. (ed.), *Progress in Experimental Personality Research.* Vol. 1. New York, Academic Press, 1964. Pp. x, 368. $9.50.

MAIER, NORMAN R. F., *Psychology in Industry* (3rd ed.). Boston, Houghton Mifflin Company, 1965. Pp. xiv, 718. $8.95.

MARRIS, PETER, *The Experience of Higher Education.* New York, The Humanities Press, Inc., 1964. Pp. v, 220. $5.00.

MCCURDY, HAROLD G., *Personality and Science: A Search for Self-Awareness.* Princeton, N.J., Van Nostrand, 1965. Pp. viii, 151. $1.45 (paper).

MERLEAU-PONTY, MAURICE (trans. by ALDEN L. FISHER), *The Structure of Behavior.* Boston, Mass., Beacon Press, 1963. Pp. xxxviii, 257. $8.50.

MONTESSORI, MARIA, *Dr. Montessori's Own Handbook.* Cambridge, Mass., Robert Bentley, 1965. Pp. xii, 121. $5.00.

MORSE, WILLIAM C.; CUTLER, RICHARD L.; and FINK, ALBERT H., *Public School Classes for the Emotionally Handicapped: A Research Analysis.* Washington, D.C., The Council for Exceptional Children, 1964. Pp. viii, 142. $2.50 (paper).

MUNN, NORMAN L., *The Evolution and Growth of Human Behavior* (2nd ed.). Boston, Houghton Mifflin Company, 1955. Pp. xi, 594. $8.50.

MYKLEBUST, HELMER R., *The Psychology of Deafness: Sensory Deprivation, Learning, and Adjustment.* (2nd ed.). New York, Grune & Stratton, 1965. Pp. xii, 423. $7.75.

OLIVEN, JOHN F., *Sexual Hygiene and Pathology: A Manual for the Physician and the Professions* (2nd ed.). Philadelphia, J. B. Lippincott Company, 1965. Pp. xiv, 621. $12.50.

ONG, JIN, *The Opposite-Form Procedure in Inventory Construction and Research.* New York, Vantage Press, 1965. Pp. 72. $2.50.

OSGOOD, CHARLES E. and MIRON, MURRAY S., *Approaches to the Study of Aphasia. A Report of an Interdisciplinary Conference on Aphasia.* Urbana, Ill., Univ. of Illinois Press, 1963. Pp. vii, 210. $5.00.

PATKA, FREDERICK, *Value and Existence: Studies in Philosophic Anthropology.* New York, Philosophical Library, Inc., 1964. Pp. vii, 239. $4.75.

PECK, ROBERT F.; HAVIGHURST, ROBERT J.; COOPER, RUTH; LILIENTHAL, JESSE; and MORE, DOUGLAS, *The Psychology of Character Development.* New York, John Wiley & Sons, Inc., 1965. (First published 1960.) Pp. xix, 267. $1.45 (paper).

PIAGET, JEAN (trans. by Marjorie Gabain), *The Moral Judgment of the Child.* New York, The Free Press, 1965. Pp. 409. $2.45 (paper).

POLATIN, PHILLIP and PHILTINE, ELLEN G., *How Psychiatry Helps.* New York, Macmillan, 1965. (First published 1949.) Pp. 288. $1.50 (paper).

POLLMER, ELIZABETH, *Alcoholic Personalities: A New Approach to the Problem of Alcoholism and its Sociological Implications.* New York, Exposition Press, 1965. Pp. 157. $4.00.

REDL, FRITZ and WINEMAN, DAVID, *Children Who Hate.* New York, The Free Press, 1965. (First published 1951.) Pp. 286. $2.45 (paper).

REDL, FRITZ and WINEMAN, DAVID, *Controls from Within: Techniques for the*

Treatment of the Aggressive Child. New York, The Free Press, 1965. (First published 1952.) Pp. 332. $1.95 (paper).

REED, C. F.; ALEXANDER, I. E.; and TOMPKINS, S. S. (eds.), *Psychopathology: A Source Book.* New York, Wiley, 1965. (First published 1958.) Pp. xii, 803. $2.95 (paper).

RICHTER, CURT PAUL, *Biological Clocks in Medicine and Psychology.* Springfield, Ill., Charles C Thomas, 1965. Pp. v, 109. $8.50.

RIOCH, DAVID MCK. and WEINSTEIN, EDWIN A. (eds.), *Disorders of Communication: Proceedings of the Association for Research in Nervous and Mental Disease, 1962.* Baltimore, Md., Williams & Wilkins, 1964. Pp. xiv, 519. $22.00.

ROE, ANNE and SIEGELMAN, MARVIN, *The Origin of Interests.* Washington, D.C., American Personnel and Guidance Association, 1965. $2.75 (paper).

ROSS, STRANGE, *Logical Foundations of Psychological Measurement.* Copenhagen, Denmark, Munksgaard, 1964. Pp. 152. 36 Kroner (paper).

ROWE, CLARENCE J., *An Outline of Psychiatry.* (4th ed.). Dubuque, Iowa, Wm. C. Brown, 1965. Pp. x, 177. $3.00.

SANDOZ, A. G. (ed.), *Psychopathology and Pictorial Expression: An International Iconographical Collection. I. Painting and Chemotherapy,* by J. Delay and R. Volmat; II. *Painting as a Means of Psychodiagnostic Research. Relationship with the Rorschach Test,* by P. L. DeLuca and B. Sacchettini; III. *The Interpretation of Visual Hallucinations,* by H. Leuner. Basel and New York, S. Karger, 1964. $11.80.

SARASON, IRWIN G. (ed.), *Psychoanalysis and the Study of Behavior.* Princeton, N.J., Van Nostrand 1965. Pp. x, 173. $1.45 (paper).

SARASON, IRWIN G. (ed.), *Science and Theory in Psychoanalysis.* Princeton, N.J., Van Nostrand, 1965. Pp. xi, 205. $1.75 (paper).

SCHILDER, PAUL, *The Image and Appearance of the Human Body: Studies in the Constructive Energies of the Psyche.* New York, John Wiley & Sons, Inc., 1965. (First published 1950.) Pp. 353. $1.65 (paper).

SCOTT, JOHN PAUL and FULLER, JOHN L., *Genetics and the Social Behavior of the Dog.* Chicago, Univ. of Chicago Press, 1965. Pp. xviii, 468. $12.50.

SKURNIK, LARRY S. and GEORGE, FRANK, *Psychology for Everyman.* Baltimore, Md., Penguin Books, 1965. Pp. 128. 95¢ (paper).

SMITH, SALLY LIBERMAN, *Nobody Said It's Easy. Can the Years Between 13 and 19 Be the Best Years of Your Child's Life?* New York, Macmillan Company, 1965. Pp. 223. $4.95.

STENGEL, ERWIN, *Suicide and Attempted Suicide.* Baltimore, Md., Penguin Books, 1965. Pp. 135. 95¢ (paper).

STRAUS, ERWIN W. (ed.), *Phenomenology: Pure and Applied. The First Lexington Conference.* Pittsburgh, Pa., Duquesne Univ. Press, 1965. Pp. vii, 208. $5.95.

TURNER, ARTHUR N. and LAWRENCE, PAUL R., *Industrial Jobs and the Worker. An Investigation of Response to Task Attributes.* Boston, Mass., Harvard Business School, 1965. Pp. xvi, 177. $4.00.

WINN, RALPH B., *Dictionary of Hypnosis.* New York, Philosophical Library, 1965. Pp. 124. $3.75.

WOLMAN, BENJAMIN B. and NAGLE, ERNEST (eds.), *Scientific Psychology: Thirty Leading Scholars Evaluate the Scientific Bases of Today's Major Psychological Systems.* New York, Basic Books, 1965. Pp. xv, 620. $12.50.

YATES, AUBREY J. (ed.), *Frustration and Conflict: Selected Readings.* Princeton, N.J., Van Nostrand, 1965. Pp. 246. $1.95 (paper).

ZETTERBERG, HANS L., *On Theory and Verification in Sociology* (3rd ed.), Totowa, New Jersey, The Bedminster Press, 1954. Pp. v, 177. $4.50.

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

VOL. LXXVIII SEPTEMBER, 1965 NO. 3

AN ANALYSIS OF THE SATIATION-FATIGUE MECHANISM OF FIGURAL AFTER-EFFECTS

By LEO GANZ, University of California at Riverside, and ROSS H. DAY, University of Sydney

In 1933, Gibson called attention to a number of effects of prolonged viewing.[1]

For example, it was reported that mildly curved lines gradually come to appear straight, and that the chromatic fringes seen through a prism gradually disappear. These phenomena were interpreted as localized normalization-effects. Even though the stimulus-input is distorted, the perceived world approximates a set of norms (*e.g.* straightness). Later, Köhler and Wallach demonstrated that some of these effects were independent of normalization.[2] They showed that, after visual exposure to a figure, a new figure was displaced in appearance away from the locus of the first. These displacements, which were called *figural after-effects (FAE)*, could be produced both in the direction of some norm and, as readily, away from that norm. Hence, some other explanation was necessary. Köhler and Wallach developed an elegant electrophysiological field theory which not only accounted for these phenomena in a qualitative sense, but also seemed to account for the depressed contrast which results from prolonged viewing, the anisotropy of visual space, the discrepancy between the veridicality of perception, and the distances

* Received for publication May 21, 1964. The experiments below were conducted in the Hunter Laboratory of Psychology, Brown University during the tenure by one of the authors of a Post-doctoral Fellowship (MF-19, 244-C2), from the United States Public Health Service. We are grateful to Lorrin A. Riggs for his encouragement and helpful suggestions.

[1] J. J. Gibson, Adaptation, after-effect and contrast in the perception of curved lines, *J. exp. Psychol.*, 16, 1933, 1-31; J. J. Gibson and M. Radner, Adaptation, after-effect and contrast in the perception of tilted lines. (1) Quantitative studies, *ibid.*, 20, 1937, 453-467; Adaptation, after-effect and contrast in the perception of tilted lines. (2) Simultaneous contrast and the areal restriction of the after-effect, *ibid.*, 553-569.

[2] Wolfgang Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. Phil. Soc.*, 88, 1944, 306-315.

of the topographic cortical projection, as well as certain illusory phenomena and their reduction with prolonged viewing.[3] The scope of the theory demonstrated the broad potentialities which an explanation of contour-displacement might have for perceptual phenomena.

Osgood and Heyer later proposed an alternative theory to account for the same facts, but relying on more orthodox neurophysiology.[4] Both theories postulated a localized state of fatigue in the nervous system. Although other theories of a descriptive type have been suggested, only Köhler and Wallach and Osgood and Heyer specify an actual mechanism. Because these theories fail to account for a number of phenomena, recent reviewers have concluded they are unsatisfactory figural after-effects;[5] but it is doubtful whether we are as yet ready to dispense with them. The present study examines one aspect of the two models: the relationship between displacement and satiation. Several studies are reported which contribute new evidence on this relationship.

Although the two theories are quite different in their hypothetical constructs, they are alike in certain formal respects. Both postulate an hypothetical state in the visual cortex with at least two independent observable manifestations. Prolonged visual inspection of an edge or contour results in a highly localized state of satiation or fatigue which interferes or resists the subsequent development of brightness-gradients in its immediate vicinity. Furthermore, when brightness-gradients are superimposed on *asymmetrical* regions of satiation-fatigue, a displacement of that contour away from the area occurs. The greater the local asymmetry, the greater the magnitude of the displacement.

In the studies to be described, contrast-reduction has been measured by examining the rise in threshold for the detection of a visual contour. Both the threshold-elevation and displacement have been measured at a number of distances from the locus of the inspection-contour. If, as customary, $\Delta I/I$ denotes the detection-threshold for a bright target on an illuminated surround, then we can let $\Delta\Delta I/I$ be the change in threshold resulting from inspection, which serves as a measure of local change in fatigue-satiation. We let *IFD* denote the interfigural distance between the contour of the inspection-figure (*I*-figure) and that of the test-figure (*T*-figure). The gradient of fatigue, as a function of interfigural distance then is given by the expression $d(\Delta I/I)/d(IFD)$, assuming continuity. Let D stand for the repulsion-displacement and K be a constant. It can be shown that both theories predict the proportionality $-d^2\ (\Delta I/I)/d\ (IFD) \cong K \cdot D$. In other words, the displacement should be greatest at

[3] Köhler and Wallach, *op. cit.*, 316-357.

[4] C. E. Osgood and A. W. Heyer, A new interpretation of the figural after-effect, *Psychol. Rev.*, 59, 1952, 98-118.

[5] P. McEwen, Figural after-effects, *Brit. J. Psychol. Monogr. Suppl.*, 31, 1958, 87-96; H. H. Spitz, The present status of the Köhler-Wallach theory of satiation, *Psychol. Bull.*, 55, 1958, 1-28; R. H. Day, R. H. Pollack, and G. N. Seagrim, Figural after-effects: A critical review, *Austral. J. Psychol.*, 11, 1959, 15-45.

the *IFD* locus where the local threshold-rise is changing most rapidly with change in *IFD.* The displacement is a repulsion, away from the satiation-maximum. The relationship is examined in a number of experiments.

Experiment I

In the first experiment, we examined the displacement of a white line induced by prior observation of a white square. We also measured the effect of that observation on the ease with which the vertical line was detected.

Apparatus. An optical device was employed which presented to *O* a homogeneously luminous background upon which a fixation-point and an inducing figure or a pair of *I*-figures could be superimposed as brightness-increments. The optical arrangements are depicted in Fig. 1. The apparatus permitted accurate lateral displacement of the *I*-figure and of the *T*-figure. The luminance of the *I*- and *T*-figures also could be changed accurately.

The stimuli as they appear (monocularly) to *O* are drawn to scale in Fig. 2. They duplicate, in size and shape, those used by Köhler and Wallach in a series of quantitative studies of displacement-effects.[6]

Observers. The *Os* were six members of the Hunter Laboratory of Psychology, graduate students and faculty, of whom half had considerable experience in making judgments of visual stimuli and half had some experience.

Procedure. During the initial trials, *E* instructed *O* with the following words:

> I will show you a pair of lines. Tell me as quickly as you can in what direction to move the *upper* line so as to align it perfectly with the lower line. I will move it in that direction. You will say, 'stop,' as soon as the lines appear in perfect alignment, that is, the upper line should look like an upward continuation of the lower line. Do not remove your fixation from the fixation-point at any time.

The presentation of the figures and the subsequent adjustment constituted a single trial. On some trials, the upper line was displaced leftward 2-4 min. of arc before presentation, and on some it was displaced to the right to the same extent. The direction and size of these initial displacements were varied irregularly. *E* kept the speed of displacement fairly constant at approximately 1 min. of arc per sec. A trial lasted from 2-5 sec.

A displacement-session began with 7 min. of dark-adaptation. *O* then viewed the fixation-point centered on the background-field and light-adapted for 3 min. Ten practice-alignment trials followed, after which five pairs of experimental trials were given. *Each* such pair began with a pre-inspection alignment. Then there were 2 min. of inspection of the rectangle, and immediately thereafter a post-inspection alignment. The pairs of pre- and post-inspection alignments were separated by 2-min. intervals. Contour-displacement is given by the difference between the pre- and post-inspection alignments. Five interfigural distances were used, one each day: 0′, 3′, 6′, 12′, and 24′. The order in which the interfigural distances were assigned on successive days was balanced over *Os*.

The sessions for detection-threshold were identical in procedure and timing in

[6] Köhler and Wallach, *op. cit.,* 300-305.

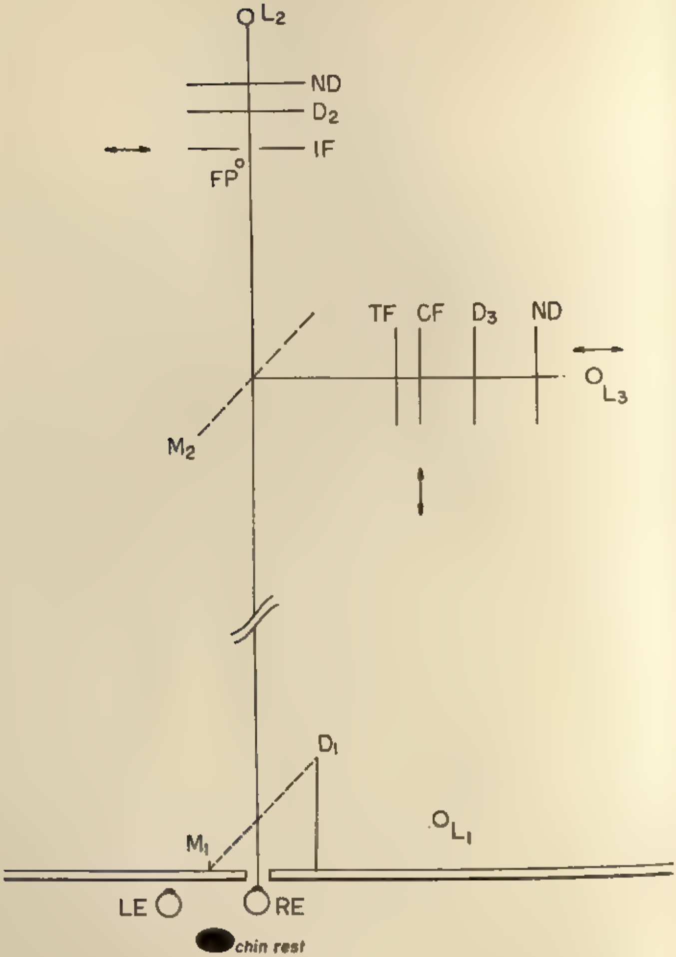

FIG. 1. SCHEMATIC REPRESENTATION OF THE APPARATUS. *O* uses his right eye (*RE*). Source L_1 is diffused by opal glass (D_1) and so reflected at M_1 as to appear as a 30° × 20° homogeneous surface (0.377 mL.) with a central fixation-point (*FP*). Source L_2, reduced in intensity (*ND*), diffused (D_2) and masked (*IF*) appears as a rectangle (5.63 mL. except in Exp. III). Source L_3 illuminates two slits, the *T*- and comparison-figures (0.944 mL. during displacement-tests). Micrometer-drives impart lateral motion to the comparison-figure (*CF*) and the *I*-figure (*IF*). A continuous gradation of the test-line luminance is effected by moving the source L_3. In the haploscopic presentation, the path originating at L_2 is presented to the left eye (*LE*). Viewing-distance is 1 m.

all respects except that the comparison-figure was omitted and the luminance of the *I*-figure was varied. Only the *T*-figure was presented. *O* was instructed initially with the following words:

[*On descending trials*] Now I will show you a single line to the right and below the fixation-point. I am going to decrease the brightness of that line and I want you to indicate by saying 'now,' as soon as you can no longer see it *It's important that you respond as soon as it disappears.* Do not remove your fixation from the fixation-point at any time.

On ascending trials, *O* was told the brightness of the line would increase, and that he would indicate by saying 'Now' when it first became visible. In other respects the timing, practice-trials, daily order of interfigural distances, and so forth, were the same as in the alignment-threshold determinations. The luminance

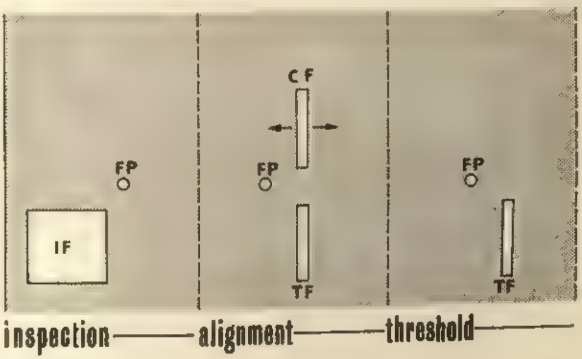

FIG. 2. STIMULUS-PRESENTATIONS DURING INSPECTION AND DURING THE ALIGNMENT- AND THRESHOLD-TESTS

of the line required to elicit the single 'Now' on ascending and descending trials for one interfigural distance was averaged over the five-trial sessions. This value divided by the luminance of the background (0.337 mL) gives the increment of the detection-threshold, $\Delta I/I$.

The use of this hybrid psychophysical method requires more explanation. Speed is important in measuring contour-displacement because of its exponential decay-function.[7] The half-life of a figural after-effect depends on brightness and on the duration of inspection, but is usually in the neighborhood of a few seconds, which limits us to a few trials at most following a single inspection-period. For speed of determination, one might turn to the method of adjustment. Köhler and Wallach, however, report a tendency on the part of some *O*s to bracket,[8] but during his bracketing *O*'s adaptation is changing rapidly. This change makes *O* less than satisfied with each bracket, and therefore he continues bracketing, while the after-effect quickly dissipates. Moreover, as the main effect diminishes rapidly, an effect from the *T*-figure itself enters the picture, all of which suggests retaining stimulus-

[7] E. R. Hammer, Temporal factors in figural after-effects, this JOURNAL, 62, 1949, 337-354.

[8] Köhler and Wallach, *op. cit.*, 300-305.

control in *E*'s hands. We have, therefore, used basically an *adjustment-technique without bracketing.*

Results. The results consist of displacements (in minutes of arc) and detection-thresholds of a *T*-line following inspection. The data for each of the six *O*s are shown in Fig. 3. We turn first to the detection-thresholds. At Point *C*, far to the right on each abscissa, is given the $\Delta I/I$ value for the pre-inspection trials. These control-trials, it will be recalled, followed at least 3 min. of light-adaptation to the homogeneous field. The log-values shown in the figure range from -0.740 to -0.595. At the other abscissa-values are given the thresholds immediately following the presentation of the *I*-figure. The effect of the *I*-figure, quite clearly, is to raise the thresholds above the control-value. With larger interfigural distances, going from 0 min. of arc (where the *I*-figure immediately adjoins the *T*-figure) to 24 min. of arc, the threshold drops monotonically. The slope of this negative function varies, though not greatly, among *O*s.

Theoretical considerations discussed above center interest on the interfigural distance (*IFD*) at which this slope is steepest. The functions depicted in Fig. 3 all appear to decelerate with increasing *IFD*. The slope is steepest very near the zero *IFD*. For 3 *O*s, LAR, JSK, and CTW, the functions appear to a first approximation-exponential, with steepest slope between 0–0.3′. *O*s FCW and RAW show very similar functions, but the greatest steepness is between 3–6′. These differences among *O*s are well within the error of measurement.

The pre-inspection alignment-value was subtracted from the post-inspection value to give a measure of displacement in minutes of arc, which yields the ordinate values for the *FAE*-curves in Fig. 3. A positive displacement indicates that during post-inspection, *O* placed the comparison-figure farther to the right than during pre-inspection. During post-inspection, presumably *O* places the comparison-figure farther to the right to obtain subjective alignment, because the *T*-figure appears farther to the right than it did prior to inspection. Since the *I*-rectangle is to the left, such a displacement to the right constitutes a *repulsion* of the *T*-contour from the inspection-figure. This is the well-known contour-displacement effect treated in detail by Köhler and Wallach. It is given in the *FAE*-curves of Fig. 3 as a positive value in minutes of arc, above the dashed line. We see from Fig. 3 that under the present conditions this repulsion occurs and has a magnitude which varies from *O* to *O*. Maximal repulsion-displacement is of the order of a 0.5-1′ in magnitude and appears at an *IFD* of 6′ or more.

Data-points below the dashed line in the *FAE*-functions of Fig. 3

(negative ordinate-values) indicate that during post-inspection O placed the comparison-figure farther to the left than during pre-inspection. Since the I-rectangle is left of the T-figure, such a move to the left constitutes

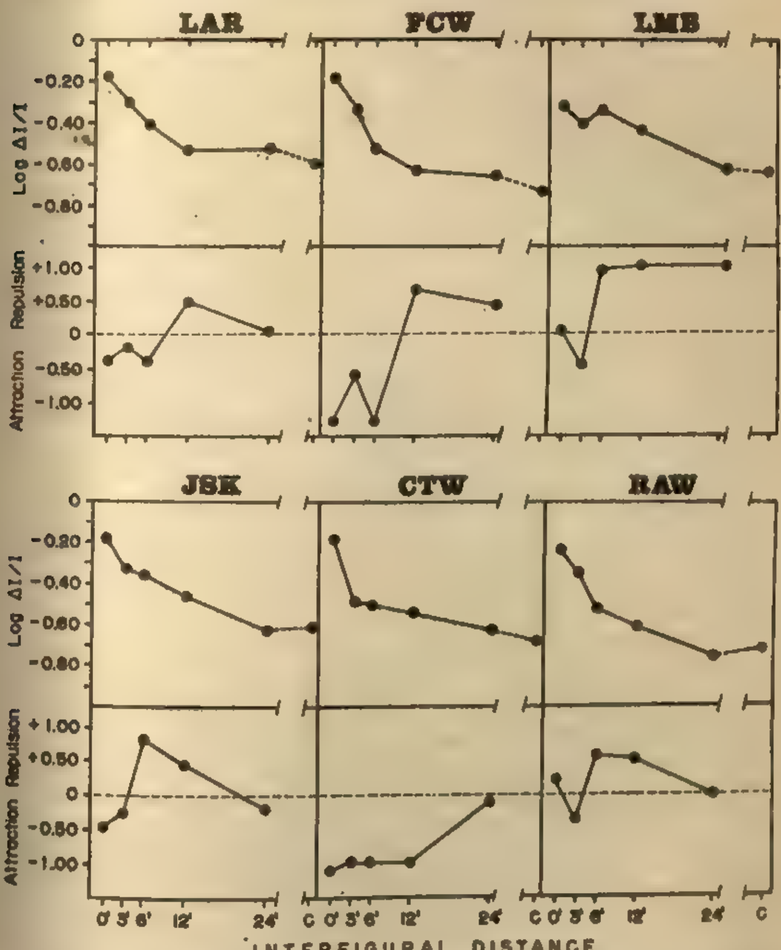

FIG. 3. THE EFFECT OF INSPECTING A BRIGHT RECTANGLE ON THE DETECTION-THRESHOLD AND ON THE ALIGNMENT OF LINES AT DIFFERENT DISTANCES FROM THE I-FIGURE CONTOUR (METHOD OF ADJUSTMENT)

an *attraction* of the T-contour to the I-figure. An attractional figural after-effect has only occasionally been reported, and thus acquires some importance in the present context. Examining Fig. 3, we see that all *O*s gave some attraction-displacements, notably at 3′ of arc *IFD*.

Experiment II

It seemed desirable, from a methodological viewpoint, to extend the generality of our findings by providing a set of converging operations. To this end, a second experiment was conducted that used a different psychophysical technique and design.

Apparatus. The apparatus was that described for Experiment I.

Observers. The *Os* were six members of the Hunter Laboratory of Psychology. They had not served in the previous experiment.

Procedure. Each *O* participated in 12 40-min. sessions, 6 displacement-sessions and 6 detection-sessions. Subjective displacement was measured by the *method of limits. O* was instructed initially as follows:

I will show you a pair of lines a bit to the right of the fixation-point for about 1 sec. Make quite sure you keep fixating the fixation-point. That is, you are not supposed to look directly at the lines. Tell me as quickly as you can how the *upper line* is located relative to the *lower* one. For 'right,' tap once; for 'left,' tap twice. The lines will reappear every 5 sec. You will each time indicate by tapping how the upper line is located relative to the lower one, for about 20 judgments. Then there will be a pause. Make sure you are fixating the fixation-point as carefully as you can.

An *L*-series of trials began with the comparison-figure some 5-10 steps to the left of the *T*-figure. *O* usually would respond by tapping twice. The comparison-figure was then moved one step to the right for the next presentation, and so on, until two consecutive single taps were given. The steps were 0.30′ of arc in size. An *R*-series of trials was identical, except that the comparison-figure was begun to the right of the *T*-figure, and so on. The series were administered at each *IFD* and were ordered LRRL or RLLR on alternate days.

A displacement-session began with 5 min. of dark-adaptation and 3 min. of light-adaptation. Six blocks of trials (0, 3, 6, 12, and 24 min. of separation and a control series without any *I*-figure) then followed to complete an experimental session. A block of trials began with a 1-min. inspection. Then the *I*-figure was removed and the following series of trials was given: (1) 0.3-sec. interval with background and fixation-point but no figures; (2) 0.7-sec. presentation of the *T*- and comparison-figures; (3) *I*-figure for 4 sec.; (4) return to Step 1; and so on. The cycling continued until the four series of trials, LRRL or RLLR, were completed. The block was completed with a 3-min. rest-period. Then a new block of trials began at a new *IFD*-setting. The four series gave four thresholds which were averaged to give a point of subjective alignment for each block. Subtracting the control-block point of subjective alignment from the five experimental blocks gave a measure of contour-displacement at each of the five *IFD*-settings in minutes of arc. This measure is comparable to the one used in Experiment I.

The detection-threshold was obtained with an identical procedure except that only the *T*-figure, not the comparison-figure, was presented. Within a block, ascending and descending series were given, either *ADDA* or *DAAD* in order, again by a method of limits and a *tap, tap-tap* indicator-response. The steps were equal log-increments (0.59 log-apparent ft-c. in size). The detection-threshold $\Delta I/I$ was estimated as in Experiment I.

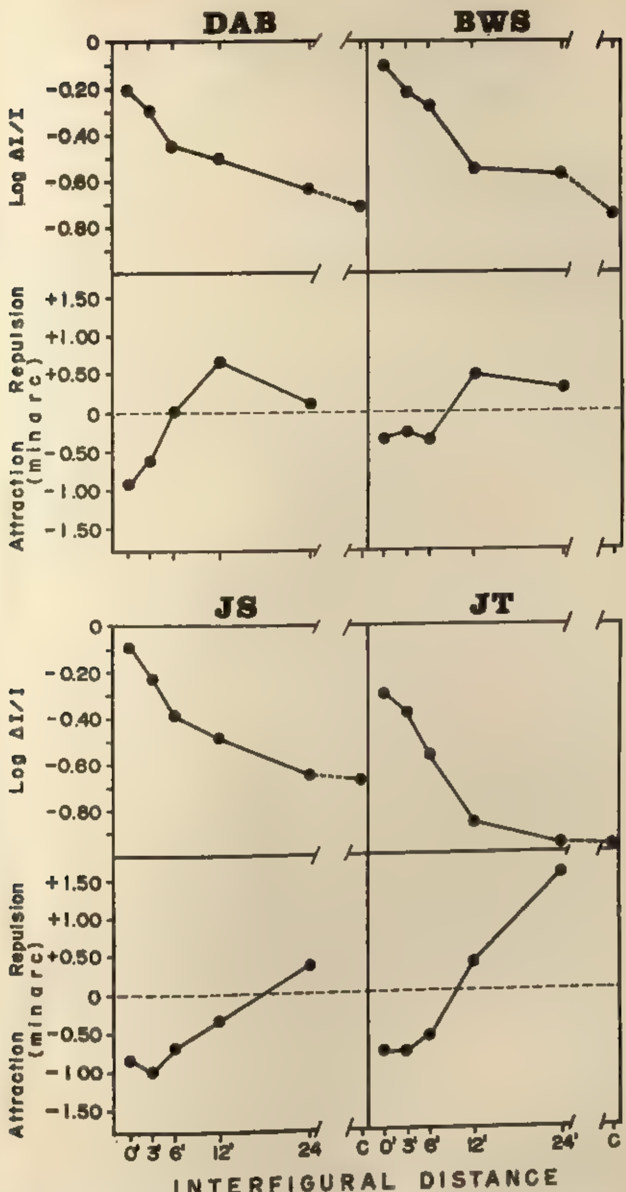

FIG. 4. THE EFFECT OF INSPECTING A BRIGHT RECTANGLE ON THE DETECTION-THRESHOLD AND ON THE ALIGNMENT OF LINES AT DIFFERENT DISTANCES FROM THE *I*-FIGURE CONTOUR (METHOD OF LIMITS)

Results. The detection-thresholds and displacement-values are shown in Fig. 4. The findings of Experiment I are duplicated in a number of basic respects. The detection-threshold is raised by presentation of the *I*-figure. When interfigural distance between the *I*-figure and the *T*-figure is larger, this rise in threshold is less. With the possible exception of those for JT, the detection-threshold functions do not appear as approximations of a Gaussian distribution. They approximate more closely an exponential function inversely proportional to the *IFD*. The steepest slope of the observed functions occurs at an *IFD* of 0-6′ of arc. If the threshold-rise function were Gaussian, the steepest slope would occur at the inflection point, *i.e.* at one *SD* from the peak or mode. Hence, 1 *SD* would be between 0-6′ of arc. At 4 *SD*s, the Gaussian function has an almost horizontal slope, and 4 *SD*s would occur well within 24′ of arc, but in all our *O*s the observed slope within 24′ descended quite noticeably. In a word, the Gaussian distribution cannot be fitted to the threshold-rise data because either the peak of that distribution is too shallow or the tail is too steep to fit our observations. The displacement-data also replicate Experiment I in a number of respects. All *O*s displaced the *T*-figure toward the *I*-figure when the interfigural distance was smaller than 6′ of arc. This is an attraction-effect once more. With larger interfigural distances, the *T*-figure was repulsed.

The following differences between Experiment I and Experiment II can be discerned: A greater rise in threshold appears to have been produced in Experiment II. For 0′, 3′, 6′, 12′, and 24′ *IFD*, the mean elevations in $\Delta I/I$ above control-levels in Experiment I were 0.370, 0.218, 0.147, 0.079, and 0.016, respectively. In Experiment II they were 0.616, 0.422, 0.298, 0.116, and 0.034. There is no overlap between these distributions until 24′ *IFD*. The greater effect in Experiment II probably is related to the smaller time-interval between *I*- and *T*-presentations. The displacement-data also differ in the two experiments. The attraction-effect is more pronounced, which is shown by the fact that a null displacement, where the function crosses the dashed line in Figs. 3 and 4, occurs at a larger *IFD* in Experiment II. Specifically, it falls at *IFD* 9, 10, 4, 3.7, and 4.2′ of arc for the *O*s of Experiment I, and 18, 10, 6, and 8.5′ of arc for the *O*s of Experiment II. It is possible that the higher threshold-rise and greater attraction-effect in Experiment II are somehow related. In Fig. 5, we see that the more the threshold has been increased at the locus of the *I*-figure contour—*IFD* = 0—the farther away the distance at which the attraction-effect gives way to the repulsion-effect. This result also suggests that higher levels of adaptation induce more attraction.

Thus, Experiments I and II suggest that the depth and extent of the attraction-effect are related to the degree to which the *I*-figure has affected the neural medium. The longer effective inspection in Experiment II and the smaller time-gap between inspection and test are associated with more pronounced attraction-effects. In turn, this means that the repulsion-effect makes its appearance and reaches maximal amplitude at a larger *IFD*. This relationship was also noted within the *O*s of the two experi-

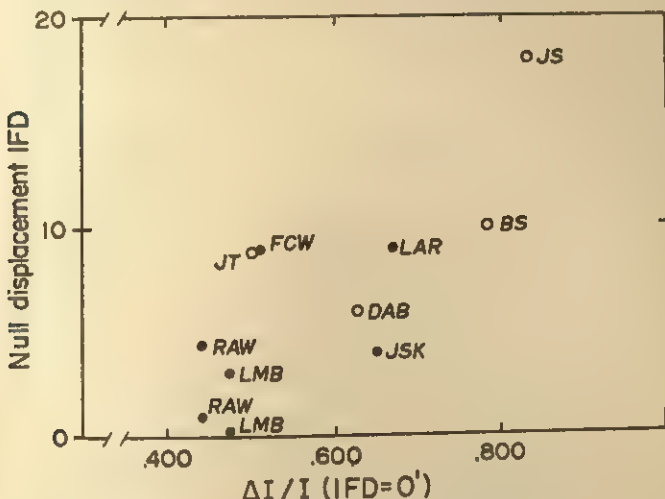

FIG. 5. THE RELATIONSHIP BETWEEN THE MAGNITUDE OF THE INCREMENT IN DETECTION-THRESHOLD, FOLLOWING INSPECTION, AND THE INTERFIGURAL DISTANCE AT WHICH THERE IS A NULL DISPLACEMENT

Solid points represent data from Experiment I, empty circles from Experiment II.

ments. The Spearman rank-difference correlation between threshold-increase and the *IFD* at which attraction and repulsion just cancel to give a 0′-displacement is +0.93.

EXPERIMENT III

The relationship between the magnitude of adaptation and the locus of attraction-repulsion was directly explored in Experiment III. The luminance of the *I*-figure was varied to obtain different strengths of effect. Inspection-duration was held constant. In accordance with the reasoning just presented, it might be expected that increasing the strength of the *I*-figure would (1) increase the distance from the inspection-contour at

which attraction is obtained, and (2) displace repulsion to a larger *IFD*. The effect of various levels of luminance of the *I*-figure on displacement at a single interfigural distance was examined.

Observers. Three graduate students of the Hunter Laboratory of Psychology served in the experiment.

Procedure. Only alignment-judgments were required in Experiment III. The *IFD* was fixed at 3′ of arc. Each *O* participated in five sessions. A session comprised 5 min. of dark-adaptation and 3 min. of light-adaptation to the homogeneous background. Six blocks of trials followed. A block began with the presentation of the *I*-figure at some luminance, after which there were four series of trials by the method of limits. The progression and timing of stimuli was

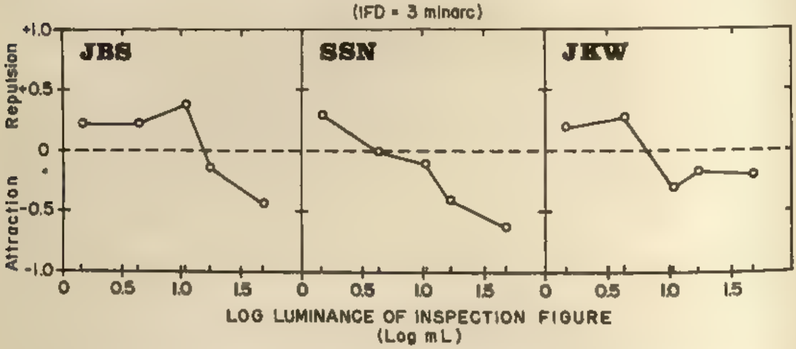

FIG. 6. THE RELATIONSHIP BETWEEN THE LUMINANCE OF THE I-RECTANGLE AND THE SIZE AND DIRECTION OF DISPLACEMENT OF THE TEST-LINE (*IFD* = 3 MIN. OF ARC.)

exactly as in Experiment II. The six blocks involved the presentation of *I*-figures at five log-luminances (0.16, 0.63, 1.02, 1.22, and 1.67 log mL.) and one control-block. There was no *I*-figure in the control-block, but only 1 min. of fixation at the fixation-point on the background-field. Blocks were separated by 3-min. intervals. The six conditions were presented in different but predetermined orders in the five successive test-sessions.

Results. The data are presented in Fig. 6 in the form of average contour-displacements. A contour-displacement, as in the experiments above, is a deviation from the setting obtained during the control condition. The data are consistent among the three *O*s in that they show repulsion with a dim *I*-figure and attraction with a bright *I*-figure.

These results are compatible with the results obtained above. A white *I*-figure of low intensity does not produce any extensive attraction. Consequently, the repulsion is seen at a small *IFD, i.e.* near the *I*-figure. As the *I*-figure gets brighter, the attraction becomes stronger. At first the

repulsion is merely attenuated. With higher luminances of the *I*-figure, a point is reached at which no displacement occurs at 3′. With still further increases in luminance, the attraction is even stronger, and an attraction-effect is manifest at the 3′ *IFD*. Thus interpreted, the results of increasing inspection-time and of reducing the interval between inspection and test in Experiment II are in some respects equivalent to increasing the *I*-figure luminance in Experiment III.

Experiment IV

A binocular presentation of the *I*- and *T*-figures offers an excellent opportunity to analyze the relationship between fatigue or satiation, on the one hand, and displacement, on the other. We know that when the *I*-figure is shown to one eye and the comparison and *T*-figures are presented to the other eye, a displacement-effect is still obtained, since this was the standard procedure employed by Köhler and Wallach.[9] We also know, partly from the same source, that some color-effects can be obtained in the same way. It has been shown, however, that light-adaptation in one eye does not raise the incremental threshold of the other eye,[10] particularly if the adaptation is long enough for a steady-state to be reached.[11] The binocular presentation thus offers a possibility of dissociating some adaptation-effects from others.

Method. Three of the *O*s used in Experiment III were employed once more. The experimental design and psychophysical method were identical to those in Experiment II. The apparatus was identical, except that the beam of the *I*-figure on the one hand and the beam of the *T*- and comparison-figures on the other hand were separated and presented haploscopically. Two fixation-points were used, one for each eye. Prior to each session, the lateral positions of these two points was adjusted until *O* fixated and fused them without difficulty.

Results. The detection-thresholds and figural displacements of the three *O*s have been plotted as solid lines in Fig. 7. The abscissa again represents the *IFD*. Dashed lines give the results of Experiment II in which the entire display was monocular. Looking first at the threshold-functions, we find the horizontal slope confirming the expectation that the threshold-rise will not appear in a haploscopic presentation. Looking next at the

---

[9] Köhler and Wallach, *op. cit.*, 271.
[10] J. F. Schearer, Effect of spatial induction on the discrimination of differences in brightness, *Phil. Mag.*, 13, 1932, 975-985; A. L. Diamond, Foveal simultaneous brightness contrast as a function of inducing- and test-field luminances, *J. exp. Psychol.*, 45, 1935, 304-314.
[11] R. M. Boynton, Some temporal factors in vision, in W. Rosenblith (ed.), *Sensory Communication*, 1961, 739-756; M. A. Bouman, On foveal and peripheral interaction in binocular vision, *Optica Acta*, 1, 1955, 177-183.

displacement-results, we see that the *T*-figure is repulsed. There is little repulsion when the *I*-figure and an edge of *T*-figure adjoin (*i.e.* when $IFD = 0$); repulsion is maximal at 12-24′ of arc. No attraction-effect is found in the haploscopic presentation. Thus, it appears that the two adaptation-effects have been dissociated.

Discussion

It seems reasonable to conclude that the repulsion of a visual contour is related to local changes in contour-sensitivity somewhere in the visual

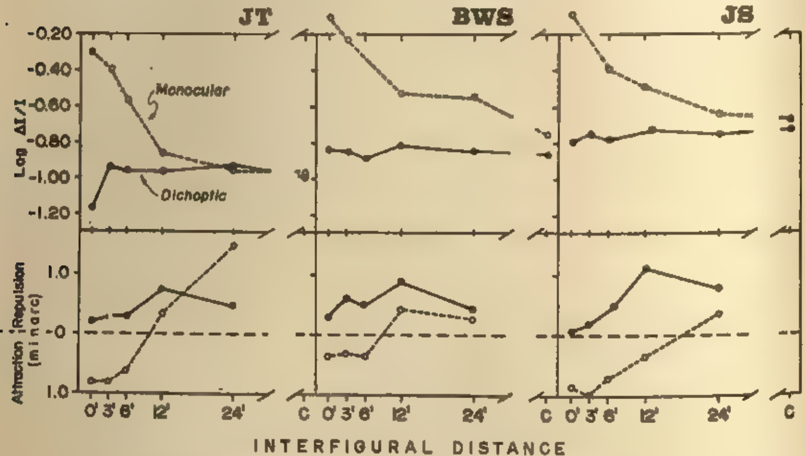

Fig. 7. A Comparison of Monocular and Dichoptic Presentations of the I- and T-Figures

system. To that extent the theories both of Köhler and Wallach and of Osgood and Heyer are correct. A white *I*-figure does raise the contour-detection threshold at the site of the *I*-figure edge. This effect, moreover, spreads beyond the immediate locus, to at least a half degree of visual angle. Higher intensities of adaptation are correlated with the changes in the locus of the repulsion, as both theories lead us to expect. Displacements are larger with higher figural contrasts.[12] We have found that intensity of adaptation is positively correlated with the spread of the dis-

[12] C. M. Freeburne and C. E. Hamilton, The effect of brightness on figural after-effect, this Journal, 62, 1949, 567-569; K. Fujiwara, and T. Obonai, Quantitative analysis of figural after-effects: II. Effects of inspection time and intensity of light stimulus upon amount of figural after-effect, *Jap. J. Psychol.*, 24, 1953, 114-120; M. Sagara and T. Oyama, Experimental studies on figural after-effects in Japan, *Psychol., Bull.*, 54, 1957, 335; Pollack, Figural after-effects: Quantitative studies of

placement-effect, but it is in the remaining quantitative details that both theories are wanting. For example, Deutsch has argued that the inflection-point of the adaptation-curve required by the Osgood-Heyer theory is at an *IFD* too large to be reconciled with the facts of visual acuity.[13] Furthermore, there are four sets of facts which have resulted from our experiments which are difficult to reconcile with either of the theories:

(*1*) It appears more plausible that the distribution of light-adaptation at distances from the inspection-contour is an exponential function of that distance, not a Gaussian function. At very small *IFD*s, this adaptation-function is changing most rapidly. Thus, the Köhler-Wallach and Osgood-Heyer theories would be forced to predict largest repulsive displacements in the region of 0-1′ of arc. This prediction is contrary to fact. Put into different words, both models must assume adaptation to have a shallow slope at 0-1′. of arc to predict the paradoxical distance-effect (maximal displacement at a non-0 *IFD*); but an independent measure of adaptation has shown this assumption is incorrect.

(*2*) The maximal repulsive displacement occurs in many *O*s at an *IFD* where the slope of the adaptation-function is quite shallow. Both models predict little or no displacement when there is near-symmetry of adaptation. These two sets of facts indicate that displacements do not follow in any simple manner (addition of tissue-resistance or subtraction of intensity) from local heterogeneities in adaptation. The predicted relationship $[-d^2\ (\Delta I/I)]/[d\ (IFD)]$, proportional to repulsion, has not been verified in these two ways.

(*3*) We obtained consistent evidence for ***attractive*** displacements at small *IFDs*. This is not a new finding, although it has been difficult to demonstrate unequivocally. Köhler and Wallach observed that when *I*- and *T*-figures nearly overlapped, the *T*-figure appeared smaller.[14] Where *I*- and *T*-figures are closed, this is an attraction-effect. Although they made light of their finding, it occurred consistently with a wide variety of figures in their studies. Similarly, Oyama has investigated displacements of a circular *I*-figure on a concentric *T*-figure.[15] When *IFD* is small, attraction often occurs. The use of concentric figures confounds attraction-effects and size-effects, but attraction-effects have been obtained also with curved lines as I-figures.[16] Dot *T*-figures, when placed near the convex side of

---

displacement, *Austral. J. Psychol.*, 10, 1958, 269-277. The effect, however, has not invariably been found (Oyama, Figural after-effects as a function of hue and brightness, *Jap. Psychol. Res.*, 2, 1960, 74-80).

[13] J. A. Deutsch, The statistical theory of figural after-effects and acuity, *Brit. J. Psychol.*, 47, 1956, 208-215.

[14] Köhler and Wallach, *op. cit.*, 271.

[15] Oyama, Experimental studies of figural after-effects: II. Spatial factors, *Jap. J. Psychol.*, 25, 1954, 195-206 (English abstract, 223-224). These findings have since been replicated (H. Ikeda and T. Obonai, Studies in figural after-effects: IV. The contrast-confluence illusion of concentric circles and the figural after-effect, *Jap. Psychol. Res.*, 2, 1955, 17-23; Y. Suto, Comparative study of assimilation-contrast illusion, figural after-effect and time error for extents of circle and line stimuli, *Jap. Psychol. Res.*, 3, 1961, 1-16).

[16] T. Yoshida, An experimental study of figural after-effect, *Jap. J. Psychol.*, 23, 1952-53, 235-238 (English abstract, 287-288); I. Kogiso, An experiment on the displacement in figural after-effects, *ibid.*, 26, 1956, 405-407.

the curve, were displaced toward the figure. This is all the more interesting because, independently, Nozawa has measured the brightness-threshold of dots with an identical stimulus.[17] He found a significant rise in threshold near the convex side and a decrease in threshold on the concave side. Thus, in a manner similar to our present findings, the *T*-figure is sometimes attracted to a locus of higher satiation. Fox, using a black rectangle as *I*-figure, found that when a black disk was placed symmetrically on the *I*-figure contour, it was, in fact, often displaced inwardly into the rectangle.[18] Smith has obtained attraction with a block-displacement not unlike Köhler and Wallach's at the large *IFD* of 30′ of arc.[19]

An attractive figural after-effect is particularly damaging to the adaptation-theories because it strikes at the core assumption: displacement of contours away from satiated or fatigued areas. That contours can be attracted to adapted areas suggests an entirely different mechanism. The present experiment indicates that attraction is present in a block-displacement proportionally identical in angular dimension to the one used by Köhler and Wallach. Why then did they not obtain attraction? There are at least two plausible reasons. First, they used binocular presentation, which we have found to weaken the attractive component. Secondly, experiments currently in progress indicate that a white *I*-figure induces attraction more readily than a black one. Köhler and Wallach used a black *I*-figure, but that presents additional problems for satiational adaptation-theories since both predict the equivalence of white-on-black and black-on-white *I*-figures.

(*4*) The displacement-effect transfers interocularly, but the adaptation-effect does not. Such a finding clearly is inconsistent with the two theories. Why should displacement occur in the absence of a measurable change in the ease with which contours are detected? Köhler and Wallach were quite aware of the possibility of dissociating color-effects, as they called them, from the displacement-effect,[20] but they did not pursue the difficulties which such facts posed for their model. Since both the color- and the displacement-effects inevitably follow from the definition of their hypothetical constructs, they should not in theory be dissociated. Why should a displacement, which implies the existence of a local heterogeneous state of anelectrotonus, occur when the same anelectrotonic state does not, under some conditions, alter the contrast (and hence the detectability) of that figure? If detection is unaffected, however, that implies no change in the readiness with which figural currents flow (in Köhler and Wallach's terms), at least at the cortical level. The argument is identical for Osgood and Heyer's model, except that it is even more explicitly a fatigue-model, and hence the possibility of dissociating fatigue and displacement is even more damaging.

We conclude, then, that there is ample evidence that contour-displacement results from local heterogeneities in the state of adaptation of the

[17] S. Nozawa, An experimental study on figural after-effect by the measurement of field strength, *Jap. J. Psychol.*, 3, 1956, 15-24.

[18] B. H. Fox, Figural after-effects: Satiation and adaptation, *J. exp. Psychol.*, 42, 1951, 317-326.

[19] Kendon Smith, 'Attraction' in figural after-effects, this Journal, 67, 1954, 174-176.

[20] Köhler and Wallach, *op. cit.*, 288.

visual system, but *displacements are not simply repulsions of figural processes from fatigued areas or areas of higher electrical resistance.*

SUMMARY

The prolonged inspection of visual figures has been reported to have the following two consequences: (1) visual objects are displaced away from the region of inspection; (2) visual objects are more difficult to detect in this region. The hypothetical constructs employed in work on figural after-effects suggest specific relationships between these two consequences of prolonged fixation. In the present experiments, the displacement and detection of test-figures at a variety of distances from the inspection-figure were measured. The results indicate that displacements sometimes occur *toward* areas of heightened detection-thresholds. Furthermore, displacements occur from one area to another which do not differ with respect to their detection-thresholds. Finally, when a dichoptic presentation is employed, displacements can be induced without concomitant threshold-changes. It is concluded that figural after-effects probably are not due to the displacement of figures away from regions of greater neural fatigue or 'satiation.'

THE CONDITIONING OF COLOR-PERCEPTION

By CELESTE McCOLLOUGH, Oberlin College

For a period of about 60 days, let *O* wear spectacles divided vertically into parts, with yellow glass on the right side and blue glass on the left side of each lens. According to Ivo Kohler, the apparent hue of a foveally fixated neutral test-field will vary for such an *O* depending upon the direction of gaze.[1] When, with spectacles removed, *O* now turns his head to the left and fixates the test-field with gaze-right, the test-field appears blue. When he turns his head to the right and fixates the same test-field with gaze-left, the test-field appears yellow. The after-effect for gaze-right is, then, the after-effect which would be expected had *O* simply worn yellow spectacles for a suitable period. The after-effect for gaze-left corresponds to the expected after-effect of blue spectacles. The novelty here is the report that the same region of the retina can show alternately the effects of blue-adaptation and of yellow-adaptation, contingent only on the position of the eyes in the head.

Kohler explains this finding as an example of "bedingte Empfindungen" (conditioned sensations).[2] "The movement of the eyes, either to right or left, seems to act as a signal for the foveal area to switch over in its color response, compensating for a yellow image in one case and a blue image in the other."[3] J. G. Taylor, whose theory of space-perception is based upon the conditioning of motor responses to visual input, accepts the concept of "conditioned sensations" as an explanation of Kohler's report of gaze-contingent color experiences.[4] According to Taylor's more explicit account, the perception of a particular color is to be understood

* Received for publication May 15, 1964. These experiments were begun under NIMH Grant MY-4056 and completed under NINDB Grant 04898-01. The author gratefully acknowledges the assistance of Dr. Günter Wyszecki and Mr. Hilton Wright, Applied Physics Laboratory of the National Research Council of Canada, Ottawa; of Dr. Dorothea Hurvich and Dr. Leo M. Hurvich, Department of Psychology, University of Pennsylvania; of Dr. R. M. Boynton, Deparmtent of Psychology, University of Rochester; and of Dr. Jo Ann Kinney and Dr. Forrest L. Dimmick, USN Medical Research Laboratory, Groton, Connecticut.

[1] Ivo Kohler, Über Aufbau und Wandlungen der Wahrnehmungswelt; insbesondere über "bedingte Empfindungen," *Sitzungsber. Öster. Akad. Wiss. (Phil.-Hist. Kl.)*, 227, 1951, 1-118; Experiments with goggles, *Sci. Amer.*, 206, 1962, 62-72.

[2] Kohler, *op. cit.*, 3ff.

[3] Kohler, *op. cit.*, *Sci. Amer.*, 206, 1962, 68.

[4] J. G. Taylor, *The Behavioral Basis of Perception*, 1962, 255-259.

as a response of the central nervous system. Such a response is normally controlled by the wavelengths of light received from the object and its surround. In this view, prolonged wearing of two-color spectacles simply brings this color-response under the additional stimulus-control of cues signaling eye-position. The results of Kohler's experiments have been widely quoted, and their importance has been underscored by Hurvich and Jameson, and by Attneave.[5] The present study was undertaken to

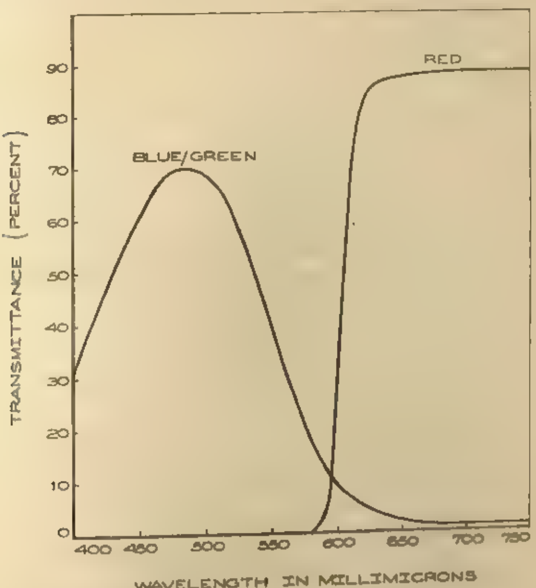

FIG. 1. SPECTRAL TRANSMISSION-CURVES FOR THE SPECTACLE FILTERS

obtain a more precise specification of the nature and extent of these gaze-contingent color changes.

Method. Following preliminary colorimetric tests, the author put on experimental spectacles and wore them continuously, except during sleep, for a period of 75 days (March 9-May 23, 1963). The colorimetric tests were repeated on Days 39 and 40 of the period of adaptation, and again on Days 66-75. Color-naming tests and standard color-vision tests also were made during the final testing period. The experimental spectacles were removed for certain of these tests, but they always were replaced as soon as the test had been completed.

---

[5] L. M. Hurvich and Dorothea Jameson, Color vision, *Ann. Rev. Psychol.*, 11, 1960, 105; Fred Attneave, Perception and related areas, in Sigmund Koch (ed.), *Psychology: A Study of a Science*, 4, 1962, 646-647.

The spectacles differed from those used by Kohler in two respects: The colors used were red and blue-green, and these colors were placed only before the right eye. The left eye was occluded by opaque plastic which completely enclosed the field of the eye. During the entire 75-day period, the left eye received stimulation for only 5-10 min. a day, while *O* was dressing in the mornings; the right eye remained closed at such times, and only the left eye was used.

Care was taken to insure that the two colored fields were used about equally by the right eye. When she was reading, *O* kept her head in such a position that the left page (left half-page, after practice) could be read through the blue-green filter and the right page (later, right half-page) through the red filter. With a combination of incandescent and fluorescent illumination, the two colored fields appeared about equally bright. In outdoor activity, *O* kept the boundary near the center of the field to which she was attending and deliberately moved her eyes in such a way as to observe objects in both colored fields.

Spectacles. The right lens contained a Bausch & Lomb Signal Green (blue-green) filter 2.0 cm. wide on the nasal side and a Bausch & Lomb Signal Red (orange-red) filter 2.5 cm. wide on the temporal side. The two filters were cemented to a single disk of clear glass and neatly joined at a vertical boundary, which lay midway between the center of the spectacle-frame and the center of the right pupil. For approximately equal transmission in daylight, the blue-green filter was made thicker than the red. Opaque plastic attached to the spectacle frames occluded the periphery of the field of the right eye and the entire field of the left eye.[6]

The two filters chosen were only approximately complementary to each other. Their hues in neutral illumination depend, of course, on the spectral composition of the light-source employed. Fig. 1 shows the spectral-transmission curves for these filters. Fig. 2 includes the color coördinates for the light transmitted by these filters, using as source the neutral mixture which served as a standard in the majority of the colorimetric measurements.

Colorimetric measurements. Since only the right eye was exposed to the adapting colors, it was anticipated that differences in color-perception would develop between the two eyes which could be measured by interocular color-matching. Most of these measurements were made on a MacAdam binocular colorimeter at the Applied Physics Laboratory of the National Research Council of Canada.[7] This instrument provides a large bipartite field, each half presenting a variable mixture of two or three color-components. With a mask placed over the exit-window, *O*, seated 30 in. from the window, saw two round spots of light 2° in diameter and approximately 2° apart. A card held in the line of sight prevented *O* from seeing more than one of the two spots at a time during the matching procedure. By manipulating this card, while opening one eye at a time, she could compare successively one spot seen by the right eye with the other spot seen by the left eye.

The right field of the colorimeter was set to give a constant daylight (neutral) mixture of luminance 8.4 ml. The left field served as a comparison-field, the color and brightness of which could be varied by *O*. When the left eye viewed the standard

[6] The spectacles were made at the R. N. Taylor Optical Company of Montreal by Mr. Albert Berg in consultation with Mr. Edward Jolley.

[7] For a general description of MacAdam's binocular colorimeter, see D. L. MacAdam, Loci of constant hue and brightness determined with various surrounding colors, *J. opt. Soc. Amer.*, 40, 1950, 589-590.

field and the right eye viewed the comparison-field, a setting of equality gave a measure of "apparent neutral" in terms of the mixture required by the right eye to match the neutral field seen by the left eye. Such determinations of apparent neutral were made with and without spectacles before, during, and after the period of adaptation to the spectacles. By a modification of the same procedure, it was possible to determine how red (or how blue-green) the neutral field appeared to the right eye when it was viewed through the red (or blue-green) filter. For such determina-

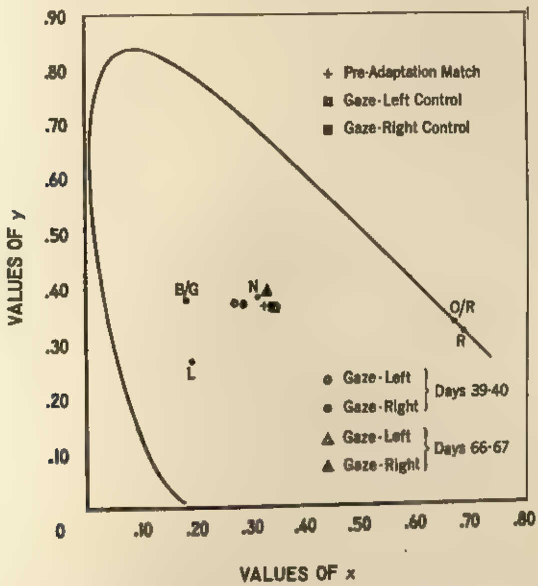

FIG. 2. MEANS FOR INTEROCULAR APPARENT-NEUTRAL MATCHES IN (x,y)-CHROMATICITY-COÖRDINATES OF THE ICI SYSTEM
(N = neutral standard field; B/G = blue-green component in colorimeter; O/R = orange-red component in colorimeter; L = left spectacle-filter with source N; R = right spectacle-filter with source N.)

tions, the right eye viewed the standard neutral field through one of these filters, and the comparison-field was adjusted until it appeared, to the left eye, exactly as red (or blue-green) as the standard field seen by the right eye.

All colorimetric matches, whether with or without spectacles, were made in duplicate series, one series with the eyes in the gaze-right position and one series with gaze-left. Observations were made in a darkened room after about 5 min. of dark-adaptation. *O* wore a pair of spectacle frames which carried, in addition to any filter to be used in the measurement-series, an opaque blind restricting right-eye vision to the right or left visual field and thereby forcing her to use gaze-right or gaze-left for the series of observations.

O controlled the mixture in the comparison-field by pressing buttons to increase or to decrease the amount of each of two components in that field. In a single session, 15 settings for gaze-left and 15 for gaze-right were made; a second session under the same conditions followed, usually on the next day, with the series of gaze-left and gaze-right settings being made in the opposite order.

A further series of 15 settings each for apparent neutral with gaze-right and with gaze-left, and a control series of 15 successive-comparison matches with left eye alone, were made on Days 74 and 75 at the USN Medical Research Laboratories. Another colorimeter of the MacAdam type was used, but in this case the comparison-field contained a mixture of three components (red, green, and blue). Except for a difference in viewing distance (18 in.), the procedure used was like that in the measurements made at Ottawa.

Color-naming experiments. On Days 72 and 73, the effect of the two-color adaptation upon color-naming was tested by a method developed at Rochester, New York.[8] Flashes of monochromatic light from an equal-energy spectrum, 300 m.sec. in duration and 3° in diameter, were presented to the dark-adapted eye in Maxwellian view. *O* used four-response keys, signaling the color names blue, green, yellow, and red; for each presentation, two responses were permitted, the first indicating the principal hue perceived and the second indicating the secondary hue, if any. A method of constants was used, with 10 presentations of each wavelength during a session. A total of 20 responses to each of 27 wavelengths from 440-600 mμ was obtained for the right eye in each gaze-direction and for the left eye in gaze-right. There also were 10 observations for each wavelength made in an exactly corresponding series with a 9° yellowish surround.

More informal measurements of color-naming were also made at the University of Pennsylvania. On Day 69, brief flashes of monochromatic light (2° in diameter against a yellowish neutral background) were shown by a modified method of limits for determination of unique blue, unique green, and unique yellow. These observations were made with each eye and in each of the two gaze-directions, at each of two luminance levels. On Day 70, four series of Munsell hues (diagonals through 5*GY* and 5*BG* at *V*3 and *V*8) were viewed in Macbeth daylight with and without spectacles, in each gaze-direction and with each eye, and the apparent colors were named.

Standard color-vision tests. Matches for color and for brightness were made with a Nagel anomaloscope without glasses for each eye and in each of three gaze-directions (right, left, and median) on Day 68.[9] On Day 71, the Farnsworth Dichotomous Test for color-blindness was attempted twice with the right eye alone, once while looking through the red filter and again while looking through the blue-green filter. These tests were repeated several months after the spectacles had been removed.

Results

Colorimetric measurements. Fig. 2 presents the results of interocular "apparent neutral" matches made on four different occasions: directly before wearing the spectacles; on Days 39 and 40 of spectacle-wearing;

[8] Measurements made by R. M. Boynton at the University of Rochester.
[9] Measurements made at the University of Pennsylvania.

on Days 66 and 67, and 8.5 mo. after removal of the spectacles. In all these measurements, the colorimeter contained two components in the comparison-field, which was viewed by the right eye without spectacles. These two components, a blue-green and an orange-red, also are indicated in Fig. 2 by the points marked B/G and O/R. All mixtures selected by the right eye had necessarily to fall on a straight line between these two points.

The point marked *N* in Fig. 2 represents the color-point calculated for the daylight-mixture viewed by the left eye and serving as a standard in the interocular matches. The mean of 30 control matches made before wearing the spectacles is indicated in Fig. 2 by a cross. Matches made on Days 39 and 40 of continuous spectacle-wearing are indicated by an open circle (mean of 30 settings with gaze-left) and a filled circle (mean of 30 settings with gaze-right). Each of the three means is significantly different from each of the others ($p < 0.001$ by t-test with 58 df).

On Days 66 and 67, it was not possible to obtain such matches with the blue-green and orange-red components alone; a small amount of yellow-green had to be added. Means for these series are represented in Fig. 2 by an open triangle (gaze-left) and a filled triangle (gaze-right); these means did not differ significantly from each other. Measurements made with a different colorimeter on Days 74 and 75 likewise showed no significant difference between gaze-left and gaze-right settings, although both series differed significantly from the control series of matches made by the left eye alone.

These three comparisons show no substantial gaze-contingent difference in color-perception. In two of the three comparisons, differences between gaze-right and gaze-left do not reach significance. In the first of the three comparisons, the difference is significant but extremely small. Two observations during the following January, 8.5 mo. after the spectacles were removed, make it impossible to regard even this small difference as evidence for conditioned color-perception. (1) In gaze-right observations, when the head was turned far to the left, the field of vision for the left eye was limited sharply by the bridge of the nose. Unless *O* shifted her head-position very slightly in changing from right-eye to left-eye observation, the standard field was viewed in the penumbra of the shadow produced by the nose. A control series of observations, made with the head deliberately held in a constant position in order that the standard field would remain in this shadow, did show a shift of its mean toward the red. Observations made on Days 39 and 40 could have been affected more than later observations by this error, since *O* became more adept in the observational technique with further practice. (2) A difference between gaze-left and gaze-right observations, significant at the 2% level, was obtained in January after 8.5 mo. of normal vision; these means are shown as open and filled squares, respectively, in Fig. 2. Such a difference cannot be attributed to conditioned perception.

Matches for "apparent neutral" made through the blue-green spectacle before spectacle-wearing, on Days 39 and 40, and on Days 66 and 67, revealed no significant difference between the gaze-right and gaze-left conditions. Means of all series made following spectacle-wearing were significantly different from the control series beyond the 1% level. No "apparent neutral" matches were possible through the nearly monochromatic red spectacle.

Results of matches for "apparent red" and "apparent blue-green" are

TABLE I

X-Coördinates for Apparent Red Matches*

Day:	62	63	63	67
Gaze-direction:	Right	Right	Left	Left
Order given:	Second	First	Second	First
Trial 1	.6852	.6797	.6932	.6779
2	.6863	.6695	.6919	.6853
3	.6836	.6740	.6968	.6969
4	.6822	.6741	.6901	.6809
5	.6897	.6850	.6864	.6819
6	.6861	.6889	.6856	.6836
7	.6837	.6799	.6982	.6748
8	.6803	.6898	.6883	.6836
9	.6766	.6897	.6859	.6833
10	.6714	.6908	.6872	.6786
11	.6738	.6869	.6840	.6837
12	.6725	.6956	.6875	.6806
13	.6745	.6830	.6754	.6878
14	.6735	.6857	.6765	.6806
15	.6726	.6952	.6907	.6832
Mean†	.6795	.6845	.6878	.6828
SD	.0056	.0075	.0057	.0044

* Color-point for red standard (right eye): $x=0.7165$, $y=0.2835$. Color-points for mixture-components (left eye): red, $x=0.7213$, $y=0.2786$; green, $x=0.2443$, $y=0.6926$.

† Gaze-right vs. gaze-left, $t=1.833$ ($p>0.10$); first vs. second series, $t=0$.

presented in Tables I and II, respectively. Since all matches were obtained as two-component mixtures, the color-points for these matches may be represented for statistical purposes by the x-coördinate alone. The x-coördinates for "apparent red" are shown in Table I in the order in which they were obtained; those for "apparent blue-green" are shown similarly in Table II. From these Tables, three findings should be noted. (1) There is no significant gaze-direction difference for either apparent red or apparent blue-green ($p > 0.10$ for right- vs. left-gaze in both tables). (2) There is no progressively increasing color-adaptation to red during a series of 15 measurements, although the right eye repeatedly viewed the neutral field through the red filter ($t = 0$ for first vs. second

series in Table I). There is, however, evidence of such a progressive increase in blue-green adaptation ($p < 0.001$ in Table II). (3) The left eye, choosing the mixture in the comparison-field to match the red (or blue-green) seen by the right eye, requires a less saturated red (or blue-green) than that actually presented to the right eye through the red (or blue-green) filter. The fovea of the right eye thus appears to have been less sensitive than the left fovea to *both* red and blue-green illumination.

TABLE II

X-Coördinates for Apparent Blue-Green Matches (Day 63)*

Session:	Morning	Morning	Afternoon	Afternoon
Gaze-direction:	Left	Right	Right	Left
Order given:	First	Second	First	Second
Trial 1	.2271	.2571	.2576	.2476
2	.2309	.2464	.2549	.2386
3	.2374	.2509	.2250	.2414
4	.2454	.2551	.2498	.2483
5	.2416	.2390	.2399	.2476
6	.2292	.2521	.2422	.2523
7	.2302	.2527	.2297	.2537
8	.2326	.2528	.2387	.2510
9	.2358	.2418	.2269	.2392
10	.2386	.2383	.2307	.2469
11	.2364	.2456	.2196	.2537
12	.2436	.2430	.2260	.2447
13	.2484	.2374	.2389	.2514
14	.2446	.2459	.2341	.2561
15	.2401	.2404	.2448	.2570
Mean†	.2375	.2466	.2372	.2486
SD	.0063	.0062	.0110	.0057

* Color-point for blue-green standard (right eye): $x=0.1798$, $y=0.3541$. Color-points for mixture-components (left eye): blue-green, $x=0.1734$, $y=0.3801$; orange-red, $x=0.6631$, $y=0.3361$.

† Gaze-right *vs.* gaze-left, $t=0.46$ ($p>0.10$); first *vs.* second series, $t=5.37$ ($p<0.001$).

This difference between the left and right eyes is the only effect of the spectacles which could be shown by colorimetry.

Color-naming and color-vision tests. The interocular difference in color-perception was not large enough to show up in the results of color-naming or in the unique-hue determinations. Comparisons of the three conditions (left eye, gaze-right; right eye, gaze-right; and right eye, gaze-left) indicated that color-names were applied to spectral colors in the same way under all conditions of observations. Efforts to uncover slight interocular differences by addition of a 9° surround and by interpolating additional test-wavelengths were unsuccessful.

Munsell-hue observations, made along diagonal sequences without *O*'s

knowledge of the hues chosen, provide scant evidence even of an interocular difference. A distinct "pinkish" overlay was reported for the first series of observations made at low saturation with the right eye (the series ranging from 5*GY*/8 through neutral to 5*P*/4), but not for later series nor for observations of the same series with the left eye. It was not possible to find any other unambiguous difference in these observations.

Anomaloscopic color-matches for each eye were within the normal range both on Day 68 and in the retest 6 mo. later. Results of the Farnsworth Dichotomous Test with and without spectacles did not differ from results obtainable on unadapted and unpracticed *O*s with normal color-vision.

Color-naming and color-vision tests thus showed only quite insubstantial interocular differences in color-perception. They revealed no trace of a gaze-contingent difference in the color-perception of either eye.

Subjective observations. During the first few weeks the spectacles were worn, the colors of the spectacles appeared to become progressively less saturated. The change occurred either wholly or primarily in the right eye. When the colored filters were placed before the left eye, objects seen through each filter seemed to be just as highly colored as they had appeared to the right eye at the beginning of the experiment. Objects viewed through a spectacle-filter alternately with right and left eyes showed a striking difference in saturation, in accord with the decreased sensitivity of the right eye shown in colorimetric measurements. This difference was evident within the first two or three weeks, and it persisted in a degree sufficiently marked for subjective perception until about two hours after the spectacles finally were removed on Day 75.

Kohler reported that there were clear differences in color-perception dependent upon gaze-direction after removal of the spectacles. In his experiment, ordinary white paper appeared yellowish when the eyes were turned to the left (where the blue filter had been) and bluish when the eyes were turned to the right (where the yellow filter had been). No such impression of gaze-contingent differences ever appeared in the present experiment. With spectacles removed, an interocular difference in color-perception was observable as a very slight pink toning of neutral colors when viewed by the right eye. This after-effect was not noticeable in viewing strongly colored fields; it was best seen in looking alternately with left and right eyes at a large expanse of wall painted a neutral hue. The after-effect did not change, however, with eye-position.

Discussion

This experiment provides no evidence of a gaze-contingent difference in color-perception. Objective color-mixing responses both with and

without spectacles showed no more difference between gaze-right and gaze-left observations after 5-9 weeks of spectacle-wearing than can be found under conditions of normal vision.

The results of color-mixing obtained here are distinctly at variance with those obtained by Kohler. When Kohler's *O* made similar color-mixing observations before wearing the spectacles, he chose the same neutral gray mixture for both gaze-directions; but upon removal of the spectacles after 60 days, he chose a combination containing more blue when he looked to the left and more yellow when he looked to the right. Kohler's results would have been replicated in the present experiment only if the match for gaze-left had deviated from the standard in one direction (toward more of the blue-green component in the right eye test-field) and the match for gaze-right had deviated in the *opposite* direction (toward more of the orange-red component). The gaze-direction difference was large enough in Kohler's experiment to place the gaze-right and gaze-left mixtures on opposite sides of the *O*'s own pretest neutral.

This evidence makes it seem unlikely that Kohler's findings were the consequence of conditioned color-perceptions. If color-perceptions could be conditioned to eye-position, some degree of difference in gaze-contingent difference should have emerged in the present study. Several possible explanations of this failure to obtain gaze-contingent color perceptions may be dismissed, as follows:

(1) Kohler's experiment was binocular, whereas the present study was monocular. It is unlikely, however, that conditioning occurs only when *both* eyes look through colored spectacles. Where care is taken to prevent conflicting use of the other eye, consistent exposure of either eye surely ought to yield conditioning. Furthermore, a series of binocular experiments by Harrington also has failed to give evidence of the conditioning of hue-perception.[10]

(2) Kohler used blue and yellow filters, whereas the present investigator (and Harrington) used red and blue-green. Yet certainly there is no *a priori* reason for expecting that the particular colors used would determine whether or not compensatory perceptions can be conditioned to the position of gaze. Moreover, Harrington's spectacles were constructed with cellophane filters provided by Kohler; these filters were considered by Kohler to have the transmission-characteristics necessary for successful conditioning.

(3) Conditioned perceptions may have been present in the right eye but masked in the interocular measurements by similar changes in the left eye. To produce such masking, however, the changes in the left eye would have had to be precisely contingent upon the gaze-direction of the right eye. Yet no gaze-contingent difference in color-perception was found with either eye alone or with both eyes togteher, either in

[10] T. L. Harrington, An experiment with colored split field glasses. Unpublished Master's thesis, University of Oregon, 1964.

ordinary subjective observation, Munsell-hue tests, anomaloscopic tests, or extensive color-naming observations.

(4) The present *O* might have been too skeptical, or insufficiently motivated, to notice differences contingent upon gaze-direction. This possibility can be rejected, since the experiment was undertaken primarily to obtain detailed measurements of gaze-contingent changes which *O* fully expected to observe.

If Kohler's results were not due to conditioning, how then are they to be explained? A careful study of his reports reveals that there are three observations which call for explanation. These three observations will be examined in turn.

(1) Desaturation of the spectacle-colors. Kohler's *O* made settings of apparent neutral with the spectacles on, both before and after the period of spectacle-wearing. When he looked to the left through the blue glass, he required an excess of yellow to obtain neutral, but this amount diminished after spectacle-wearing. When he looked to the right through the yellow glass, he required an excess of blue, which also decreased after spectacle-wearing.

As Tables I and II show, these results are confirmed by the present experiment, *but they do not in themselves provide evidence for gaze-contingent color-perceptions.* The present experiment shows, in addition, that the measured desaturation is present *for each color for both directions of gaze,* and in the same amount. The measurements indicate, therefore, that the foveal region used for color-matches had undergone desensitization both to red and to blue-green.

Common beliefs about color-adaptation apparently lead many persons to expect that a local region of the eye can be adapted either to red or to its complement, but not to both at once. Such an expectation is not justified. To determine what kind of color adaptation should be expected under the conditions of this experiment, let us first consider a simpler case, that of an *O* viewing through a colored filter a source of white light which contains equal energy at all wavelengths. Exposure through the red filter will reduce the sensitivity of the eye to those wavelengths which pass that filter; exposure through the blue-green filter will decrease sensitivity to the wavelengths passing that filter. In both cases, the decrease for any wavelength will be related to the amount of light at that wavelength which the filter admits. Rapid alternation of light through the red filter with light through the blue-green filter should, therefore, produce some decrease in sensitivity at all those wavelengths passing either one filter or the other, and the extent of the decrease at any wavelength will depend on the amount of light of that wavelength which reaches the eye. *O* will see the hue which results from a mixture of light passing the red filter with light passing the blue-green filter, and he presumably will show adaptation-effects comparable to those produced by a steady light having the hue of the mixture.

It is difficult to state precisely what radiations the right eye received through the spectacle in the actual experiment, since the precise radiations depend upon the light-source and the objects viewed; but taking equal-energy white as an approximation, and weighting the spectacle transmissions by the function of visibility, one obtains an exposure-curve with a substantial peak in the yellow-green and a secondary peak in the red. The color-matches made without spectacles did indeed shift toward the green for both directions of gaze. The data are consistent, therefore, with the

interpretation that alternate exposures of the right fovea to red and blue-green produced a net effect of green-adaptation, regardless of gaze-direction.[11]

Kohler does not report data for neutral settings *with the spectacles reversed.* According to the present findings, his adapted *O*, seeking a neutral hue while looking through the blue filter, would have obtained a diminished predominance of yellow in the mixture even when the blue filter was viewed with gaze-right. Yet the hypothesis of "conditioning" requires that the conditioned mechanism always "add blue" in response to the gaze-right cue, and Kohler should predict, therefore, that *O* would select *more* yellow after spectacle-wearing than before, when the test-field is viewed through the blue filter in the reversed (rightward) position. Evidence that the blue filter appeared more blue in gaze-right and less blue in gaze-left would support the existence of "conditioned color perception." Kohler offers no such evidence, and the present experiment yielded, instead, the result which would be expected on the basis of simple retinal color-adaptation.

(2) The "candle" effect. After removing the spectacles, Kohler's *O* saw a visual field toned in yellow at the left "as if someone were walking alongside me, carrying a candle," and in "colder hues" at the right. To this *O*, the after-effect of hue in the left visual field must have been clearly distinguishable from the hue in the right visual field. Since *O* compared what he saw to an effect of colored illumination, the "candle" effect must have involved a considerable region of the visual field.

This observation does not provide evidence of conditioned color-perception. A careful analysis of the conditions imposed by Kohler's spectacles will show that the "candle" effect is predictable from color-adaptation developed in the nasal retinal fields.

Kohler's spectacles were 5 cm. in diameter, placed 2.5 cm. from the cornea. The dividing lines between blue (left) and yellow (right) were located slightly nasal to the median plane through the pupil in order to make the blue-yellow borders coincide in binocular viewing of objects at a distance of 2 to 3 m. Consider first the exposure-conditions for the left eye. When the left eye fixates an object near the nasal border of its visual field, that object and all objects within about 30° to the left are seen through the yellow filter. Objects lying more than 30° to the left of fixation are seen through the blue filter. Since the visual field to the left of fixation stimulates the nasal side of the left retina, it is apparent that the nasal retina beyond 30° will be exposed exclusively to blue under these conditions. As the left eye rotates to the left, the region of blue-exposure spreads toward the fovea and beyond. Yellow-exposure can reach more than 30° nasally in this retina only when

[11] There are other differences between the spectacle-experiment and the simple-alternation experiment described above, revealing certain gaps in our information about color-adaptation. In the spectacle-experiment, alternation took place at a much slower rate, allowing time for some recovery between successive phases of red-exposure, but the total experiment also was prolonged far beyond the time-intervals ordinarily employed in color-adaptation experiments. Brief exposures to color at ordinary intensities have rather transient after-effects (on the order of seconds). The results shown in Fig. 2 and in Tables I and II give ample evidence, however, that periods of prolonged color-adaptation must produce changes in sensitivity whose duration is of the order of minutes or hours. It should not be assumed without explicit evidence that the cumulative adaptation obtained in prolonged experiments with colored spectacles involves no processes not already at work during brief exposures to single colors.

the left eye rotates yet farther to the right; it will do so only when the right eye fixates objects which lie farther to the right, outside the region of binocular overlap.

If it may be assumed that illumination in Kohler's experiment was approximately white, and that there was no distinct bias in the spectral reflectance of the objects being viewed, the exposure-conditions of the left nasal retina may be summarized as follows: (a) Viewing of objects within the region of binocular overlap at any viewing distance permitted yellow-exposure to occur as far as 30° into the left nasal retina. Beyond 30°, only blue-exposure occurred during binocular fixation. (b) Viewing of objects visible to the right eye alone could produce yellow-exposure farther into the nasal retina of the left eye out to a maximal excentricity of 50° when the right eye fixated objects at the border of its spectacle. Assuming that vision was confined most of the time to the region of binocular overlap, the nasal retina of the left eye after adaptation may be approximately described as consisting of three vertical bands differing in their expected color-adaptation. The first band, 30° wide and extending inwardly from the fovea, would have become adapted approximately equally to yellow and to blue. An adjoining band, 20° wide, would have become adapted principally to blue and somewhat to yellow. The remainder of the color-sensitive nasal retina (beyond 50°) would have become adapted exclusively to blue.

Since some color-sensitivity to blue is present as far as 65° into the nasal retina,[12] the "candle" effect reported by Kohler's *O* can be explained as an after-effect of predominant blue-adaptation in the nasal retina of his left eye, giving a yellowish appearance to objects seen in the left periphery beyond about 30°-50° left of fixation. The "colder hues" on the right can be explained similarly in terms of the predominant yellow-adaptation of the nasal retina of the right eye. A comparable effect in the present experiment, if it had occurred, would have appeared as a blue-green tinge in the right visual periphery consequent upon predominant red-adaptation of the right nasal retina. Since the region of exclusive red-adaptation under the conditions of this experiment would have begun about 46° from the fovea, the absence of a visible after-effect is understandable, for red-green sensitivity does not extend as far as 46° from the fovea either in the nasal or the temporal retina.

(3) Gaze-contingent differences in color-mixing experiments. Kohler's *O* selected (with spectacles off) mixtures of yellow and blue which differed measurably as a function of gaze-direction. In the present experiment, as well as in Harrington's replication, no such gaze-contingent differences were present. If the gaze-contingent differences found by Kohler do not represent "conditioned color perceptions," what factor in his measuring situation could have produced them?

Dorothea Hurvich has suggested that one should not forget the relatively homogeneous adaptation of the visual periphery in these experiments, when one is considering the probable apparent color of a foveal test-object.[13] The fovea itself may be equally adapted to both blue and yellow, but the peripheral retina clearly is not. Indeed, as has been shown above, the left periphery of vision is expected to show the effect of homogeneous blue-adaptation to such an extent that a neutral surface lying as much as 30°-50° left of fixation will appear distinctly yellow. In the presence of a neutral surround, therefore, the foveally-fixated test-field may well be

---

[12] Committee on Colorimetry of the Optical Society of America, *The Science of Color*, 1953, Fig. 20, p. 104.
[13] Personal communication.

influenced in gaze-right by the *left* (yellowish) peripheral field and in gaze-left by the *right* (bluish) peripheral field.

The present *O*, and Harrington as well, used a dark surround in which no effects of homogeneous peripheral adaptation could be seen; both failed to find differences contingent upon direction of gaze. Things were otherwise in Kohler's experiment. In response to the author's inquiry about his surround-conditions, Kohler writes, "The Scheffler Farbwähler was used for measurement of the adaptation. The test-field (of ground glass, continuously variable from yellow through gray to blue) was about 5 × 7 cm., the distance from *O's* eye being about 35 cm. The immediate surround of the test-field was dark (within a binocular visual angle of about 40°). Farther toward the periphery, the visual field was illuminated. The illumination was normal daylight. It was midday at the time of the measurement. A second series was done in the evening after dark, with normal incandescent light (in the periphery)."[14]

With this information, it is possible to determine whether the surround-conditions in Kohler's experiment can have been responsible for gaze-contingent differences which he obtained. Kohler's *O* always wore spectacle-frames during the measurements. Fig. 3 shows the approximate field of view inside the spectacle-frames for each eye in each of the two gaze-directions. Fixation was at the center of a test-field approximatly 6° wide, lying within a dark region 40° wide. As shown above, the retina of each eye may be considered to have been about equally adapted to yellow and blue within a band extending approximately 30° to either side of fixation.[15] Beyond 30° in the left field of each retina, the adaptation to yellow predominated, giving a bluish appearance to neutrally illuminated objects in the right visual field. Beyond 30° in the right field of each retina, the adaptation to blue predominated, giving a yellowish appearance to neutrally illuminated objects in the left visual field. In Fig. 3, dotted lines indicate the 30° boundaries, and the letters *Y* and *B* indicate that adaptation to yellow or blue, respectively, predominated beyond those boundaries in the visual field.

If one considers those parts of the illuminated surround outside the 30° boundaries, Fig. 3 shows that only the right part was visible in the gaze-left condition and only the left part in the gaze-right condition. *A change in gaze-direction thus changed the side on which this surround was seen, thereby changing the apparent hue of the surround.* While a distance of 30° of visual angle between test- and inducing fields is surely too great a distance for ordinary simultaneous color-contrast, it may not be too great a distance for other field-effects. The apparent hue of a highly desaturated test-field is notoriously dependent upon the hue of other parts of the visual field. A test-field which is gray in the absence of other light will appear yellowish in the presence of a bluish spot of light and bluish in the presence of a yellowish spot. To obtain a report of "neutral," more blue will have to be added to the test-field when the nearby spot is bluish, and more yellow when the spot is yellowish. These shifts are precisely the ones reported by Kohler.

---

[14] Translated by the author from a personal letter. The Scheffler Farbwähler is a colorimeter constructed for Kohler's laboratory by P. Scheffler.

[15] In fact, a predominance of yellow (or blue) begins at about 23° in the temporal retina of the left (or right) eye. This nasal-temporal difference may, however, be neglected for the sake of simplicity; its consideration would merely add strength to the argument to follow.

To show that Hurvich's suggestion is correct, and that Kohler's gaze-contingent differences are due to peripheral field-effects, only one thing is lacking—proof that the magnitude of the gaze-contingent effect is not greater than the magnitude of such a field-effect. Unfortunately, we have no quantitative evidence on the colorimetric magnitude of either effect. The general size of the gaze-contingent effect can be

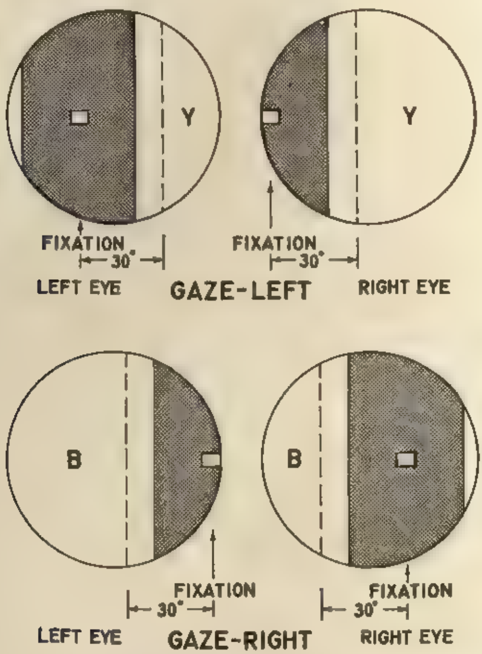

FIG. 3. APPEARANCE OF KOHLER'S TEST-FIELD AND SURROUND IN GAZE-LEFT AND GAZE-RIGHT OBSERVATIONS

(Y = adaptation to yellow, giving a bluish appearance; B = adaptation to blue, giving a yellowish appearance.)

estimated from some of Kohler's unpublished descriptions, which indicate that the magnitude of the gaze-contingent effect was quite small.[18] (a) Kohler describes the variations of hue as lying in a "very small middle region of weakest yellow- or blue-toning." The region was in fact so small that he considered it necessary to deviate from the standard psychophysical procedure for the method of adjustment, in which settings are made starting alternately from distinct blue and from distinct yellow; he started his *O* always from a neutral gray. (b) A sheet of white paper,

[18] The following observations were reported by Kohler in letters to the author.

placed on a table under overhead illumination, was seen after adaptation as yellowish with gaze-left and bluish with gaze-right. Kohler held a sheet of blue (or yellow) paper in a plane perpendicular to the table and parallel to the upper edge of the white paper. *O* was prevented from seeing this colored paper directly. By tilting the colored paper backward from the perpendicular, Kohler varied the amount of blue (or yellow) reflected from it to the white paper. The color so obtained by reflection was sufficient to counteract the gaze-contingent effect and make the white paper appear neutral. (c) The magnitude of the effect has also been described in terms of the difference between two types of ordinary writing-paper, one of which appears a "slightly soiled white," the other as having a "faint bluish tinge."

All three of these observations suggest that the magnitude of the gaze-contingent effect is much smaller than most persons seem to have assumed from their reading of Kohler's published reports. It is seen as a slight variation in the toning of a very desaturated test-field. Exploratory experiments undertaken recently in the Oberlin laboratory indicate that colors induced by chromatic adaptation and seen in a neutrally illuminated region 20° away from such a test-field can in fact shift its hue in the expected direction.[17]

The relevant quantitative experiments on such field-effects have not, however, yet been done. There is no published evidence quantifying the effect of colored stimuli in the periphery upon the hue of foveally-fixated test-fields. Nor is there evidence on the relative effectiveness of colored after-effects seen monocularly and binocularly when they are used as inducing fields for test-fields of any size or location. Such experiments are needed and overdue. If Kohler's imaginative experiments with two-colored spectacles do not after all establish the existence of conditioned color-perceptions, they may nevertheless stimulate the investigation of some neglected problems in simultaneous and successive color-induction.

Summary

For a period of 75 days, the author wore before the right eye a spectacle containing blue-green glass on the left and red glass on the right. The left eye was occluded. Objective measurements after 5 and after 9-10 weeks failed to give evidence of gaze-contingent differences in color-perception by either eye, although the right fovea showed less sensitivity than the left fovea to both red and to blue-green.

This study was undertaken as a replication of experiments reported by Ivo Kohler, in which gaze-contingent color-differences were observed following a two-month exposure to blue and yellow spectacles. Kohler

[17] Rachel Doody assisted in these experiments.

interpreted these differences as "conditioned sensations." Since the replication revealed no gaze-contingent differences, it is doubtful that Kohler's finding can be attributed to conditioning. It is suggested that homogeneous adaptation of the nasal retina to blue on one side and to yellow on the other may have played an important role in his observations. In Kohler's measurements, a neutrally illuminated surround was visible either in the left or in the right periphery, beginning at 20° from fixation. Since gaze-direction determined the side on which this periphery was visible, the apparent color of the surround likewise was dependent on gaze-direction and may have influenced the color-appearance of the foveally fixated test-field.

FURTHER OBSERVATIONS ON THE NATURE OF EFFECTIVE BINOCULAR DISPARITIES

By LLOYD KAUFMAN and COLIN PITBLADO, Sperry Rand Research Center, Sudbury, Massachusetts

In a recent paper, it was concluded that a vital question for the theory of stereopsis is "Disparity in what?"[1] An answer to this question is of importance because it would define the complexity of the phenomena with which theory must deal. We therefore set out to describe more precisely the stimulus-conditions necessary to the occurrence of stereopsis. This paper presents some of the findings of continuing research in this area.

It is clear that a number of different stimulus-dimensions can produce stereopsis.[2] Apart from disparity between similar spatial contours or forms, depth-perception occurs with disparity between similar groupings of different and rivalrous forms. It also occurs with disparity between patterns of brightness carried by similar and non-disparate forms. Brightness-disparity is also capable of eliciting the depth-response when carried by utterly different forms. The present observations suggest that all of the dimensions of stimulation-studied thus far are reducible to special instances of relative brightness-disparity. The visual processes involved in detecting such disparity are, however, more complicated than it appears on first analysis.

A simple instance of brightness-disparity is shown in Fig. 1. The letters in both patterns are alike and non-disparate, but the inner letters of both patterns are modulated in brightness to produce patterns of gray and black. Both brightness-patterns are alike except that the pattern in one field is shifted one letter-space toward one side relative to the other

* Received for prior publication April 12, 1965. This Research was partially supported by the U. S. Army Medical Research and Development Command under Contract No. DA-49-193-MD-2654. The authors are indebted to George Harker for advice and for the replication of some of the demonstrations in his laboratory. We also thank Julian E. Hochberg for many useful discussions of the problems associated with this research.

[1] Lloyd Kaufman, Some new stereoscopic phenomena and their implications for the theory of stereopsis, this JOURNAL, 78, 1965, 1-20.

[2] Kaufman, On the nature of binocular disparity, this JOURNAL, 77, 1964, 393-402.

field. Despite the correspondence of similar contours, the disparate brightness-pattern can produce stereopsis.[3]

The importance of this pattern lies in the fact that it contains no contour-disparities. Helmholtz has shown that contour is capable of eliciting a stereoscopic response in seeming independence of brightness-values.[4] His famous stereogram, in which white lines on a black ground are viewed together with black lines on a white ground, is cited frequently as evidence for the effectiveness of contour-disparities. Brightness-differences are ignored in Helmholtz's pattern, while contour-relations are ignored in the present pattern. This observation led to the following demonstra-

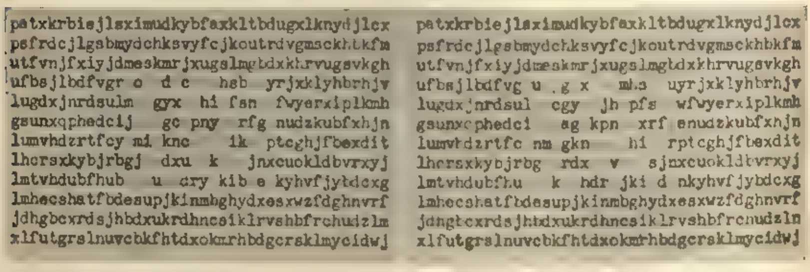

FIG. 1.

tion in which brightness-disparity was pitted against form- or contour-disparity.

DEMONSTRATION I

The same letters are employed in Fig. 2 as in Fig. 1, but the principle of construction is somewhat different. First, a contour-disparity was produced by shifting an inner region of letters in the left half of the stereogram one letter-space to the left. This operation alone produces an uncrossed disparity for the inner region relative to the background, and viewing the pattern stereoscopically therefore results in the perception of an inner rectangle behind the background. In addition to this uncrossed disparity, we superposed a crossed brightness-disparity by modulating the brightnesses of letters on the right side to produce a brightness-pattern. The same pattern was reproduced on the left side, except that it was shifted

[3] There are individual differences in the description of this phenomenon. Some *Os* report depth for the inner square as a whole, while others report depth of a veil-like surface with the letters behind it. Intermediate judgments also occur sometimes.

[4] Hermann von Helmholtz, *Handbook of Physiological Optics*, J.P.C.S. Southall (ed.), 3, 1925, 512 and Plate IV, Fig. Q.

one letter-space to the right. This operation alone would cause the inner rectangle to be seen in front of the background-letters.

The first impression reported by four *O*s who viewed this stereogram was that the inner rectangle was in front of the background. This response was to the brightness-disparity. Subsequently, for the same *O*s, the inner rectangle went behind the background. Ultimately, they found themselves capable of reversing the direction of depth at will. The phenomenon is very much like that experienced when viewing a Necker cube or an ambiguous figure-ground pattern. It was necessary to try several different levels of contrast before an easily reversible pattern was found. We are now trying to measure the amount of contrast associated with reversibility.

There is an interesting adjunct to this demonstration. We noted that

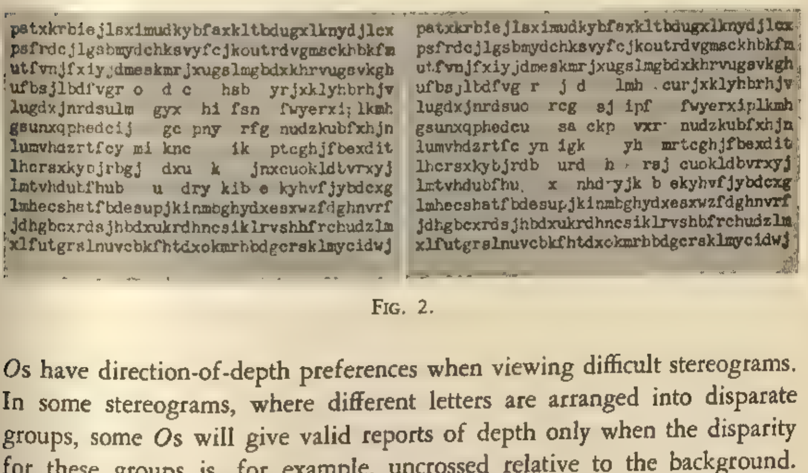

FIG. 2.

*O*s have direction-of-depth preferences when viewing difficult stereograms. In some stereograms, where different letters are arranged into disparate groups, some *O*s will give valid reports of depth only when the disparity for these groups is, for example, uncrossed relative to the background. Reports of the direction of depth will be correct when the pattern contains an uncrossed disparate inner region on a given experimental trial. In trials on which the disparity is crossed, the correctness of response remains at the chance-level. Now, when an *O* prefers a definite direction of depth for the enclosed region and he is shown Fig. 2 with the cross of the brightness-disparity consistent with his preference, he may be unable to respond to the form-disparity. When, however, the cross of the form-disparity is consistent with his preference, he can respond to either the form-disparity or the brightness-disparity at will. We therefore suggest that if the interested reader finds it difficult to observe this reversal, he should cut out the pattern and interchange the two halves. The inter-

change may tend to balance the asymmetry in response and make the observation possible.

This demonstration suggests that there are two classes of disparity capable of eliciting a depth-response, *i.e.* form-disparity and brightness-disparity, but, a different set of assumptions can explain the phenomenon equally well. It is a fact that the visual system responds to relative brightness-disparities as it does to absolute brightness-disparities. This property of the system can be demonstrated by holding a fairly dense filter over one eye and viewing Fig. 1. Depth occurs, even though the brightnesses of elements in the two fields are different. It is clear from this fact that it is not necessary that contrast-ratios in the two patterns be preserved. As long as the signs of the brightness-difference between the elements and their backgrounds are the same, stereopsis may result.[5]

If the foregoing is correct, the form-similarity in the two half-fields of Fig. 2 may be related only accidentally to the reversible-depth effect. If the visual system responds to disparate regions or patches whose only common characteristic is that they are both either more or less bright than their immediately adjacent backgrounds, the form-effect can be explained in terms of brightness-disparity. In the one case, the response is to the disparity between letters or parts of letters whose brightnesses are less than their immediate white backgrounds, whether the individual letters be gray or black. This fact would explain the so-called form-effect. In the other case, the disparity is between the brightness-relations among the letters themselves; *i.e.* black letters occupy different positions in the over-all patterns relative to the gray letters. This disparity, defined in terms of brightness-relations among letters themselves, would be an effect of a different order, involving the same stimulus-dimension as in the first case. The fact that there are sequences of gray and black letters suggests that the average areal brightness of groups of letter relative to each other is the relevant stimulus. Since both letters and groups can be seen and attended to as different aspects of the same pattern, it is not beyond the realm of possibility that these same aspects can evoke stereopsis alternatively.

It is very difficult to state the foregoing argument clearly, and for this reason, perhaps, it is not an appealing explanation. Moreover, we are still left with Helmholtz's pattern which seems to support the converse notion, *i.e.* disparities between similar spatial contours are sufficient to produce stereopsis, but Whittle has suggested a way in which the Helm-

[5] Anne Treisman (Binocular rivalry and stereoscopic depth perception, *Quart. J. exp. Psychol.*, 14, 1962, 23-27), draws this conclusion from her research.

holtz-effect can be explained in terms of brightness-disparity.[6] He proposed that, since lines have a finite width, the edge of a line in one eye could be interacting with the opposite edge of a line in the other eye, which would cause regions with similar directions of brightness-change to interact in producing the depth-effect. The following demonstration provides a test of this hypothesis.

DEMONSTRATION II

A simple Wheatstone-type stereogram is shown in Fig. 3. The gray bars on the left side are farther apart than the gray bars on the right

FIG. 3.

FIG. 4.

side. The bars are on backgrounds of opposite color. As expected from Helmholtz's demonstration, a depth-effect results when the patterns are superposed. Now, as Whittle suggests, the effect may be due to the disparity between the left sides of the bars in the left field and the right sides of the bars in the right field, thus providing two regions having identical signs of brightness-change. The stereogram in Fig. 4 shows how this explanation can be tested. There we have gray bars divided by a black region in one eye and by a white region in the other. Furthermore, the left side has a black outer background and the right side a white outer background. There are no regions in the displays which contain consistent directions of brightness-change. The writers have viewed this pattern for extended periods of time and have been unable to achieve a

[6] Paul Whittle, Binocular rivalry, Doctoral dissertation, University of Cambridge, 1963, 50-51.

depth-effect. Although one must be wary of negative results in this field, it seems clear that contour-alone is not a sufficient condition. This result is consistent with the hypothesis that the sign of a brightness-difference in adjacent regions will at least affect the ease of occurrence of a depth-response and probably is essential to it.

The foregoing result also permits an explanation of the loss of stereopsis in the Julesz-pattern in which one half-field is the negative of the other.[7] Consider a letter analogue to the Julesz-pattern in which one side of the stereogram has white letters on a black background and black letters on a white background. If the white regions in one eye interact with the lighter regions in the other eye, all of the regions of similar brightness

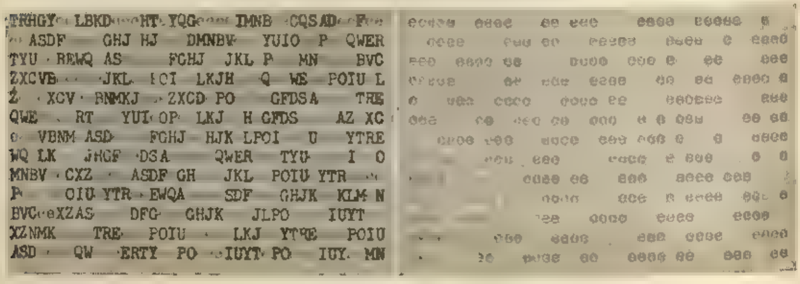

FIG. 5.

are non-disparate and should not go into depth. If form produces the response, depth should occur. Since stereopsis does fail, the result is consistent with the brightness-disparity notion.

DEMONSTRATION III

We have stated that disparities between regions having similar relative areal brightnesses are the effective stimuli for stereopsis. The following demonstration shows how this factor might operate independently of both absolute-brightness and physical-form disparities. In Fig. 5 there is a collection of black upper-case letters and gray *e*s on one side of the stereogram. The other side contains gray *e*s only, but they are arranged in groups which have the same locations and numbers of letters as do the upper-case letters on the left. Now the *e*s in both half-fields have common brightness and shape which would interact to vitiate any depth-effect if either form or absolute brightness were the effective stimulus. Since, however, the *e*s on the right are darker than their background, and since

[7] Bela Julesz, Stereopsis and binocular rivalry of contours, *J. opt. Soc. Amer.*, 53, 1963, 994-999.

the upper-case letters are darker than their background of *e*s, disparity between these regions having related areal brightness should produce stereopsis. With a little patience, the reader can see that this is indeed the case when he views the pattern in a stereoscope. It might be argued, of course, that the groupings of upper-case letters are of the same size as the groupings of *e*s in the contralateral field, and, since both have figure-properties—in the sense of figure-ground—stereopsis results. The reaction here is not due as much to brightness disparities as to a grouping disparity.

To control for the possible effect of group-size as the critical feature in the last experiment, the pattern illustrated in Fig. 6 was constructed. There, all groups of letters have been made the same size; and it will be observed at once that, if group-size were the significant factor, this cue,

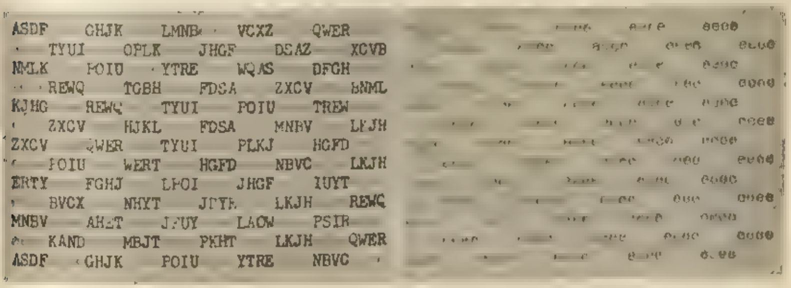

FIG. 6.

supported by identical forms and brightnesses, should cause the *e*s to interact, thereby destroying the depth-effect. In fact, however, the *e*s interact with the upper-case letters to produce stereopsis when this pattern is viewed in a stereoscope. The perceived direction of depth is consistent with the cross of the disparity between the inner *es* and the inner upper-case letters relative to the background-letters.

DEMONSTRATION IV

Although the depth-effects in Fig. 5 and 6 may not be attributed to the sizes of groups used, nor to the form or absolute brightness of the individual letters in the groups, the effect of grouping *per se* has not been isolated here and therefore cannot be ruled out completely. In an earlier paper it was reported that triadic structures of different letters, when disparately represented, can produce stereopsis, and was argued that the similar groupings are sufficient stimuli to depth. A further illustration

of the capability of groupings of dissimilar elements to produce stereopsis is provided by Fig. 7.

Entirely dissimilar sets of letters are arranged into groups of similar size, and a disparity is provided by shifting to one side by one letter space some of the groups in the right half of the stereogram. Depth results in the face of rivalry between the *l*s and *o*s in the central regions of both half-fields. That this effect is quite universal and independent of the letters used is shown by Fig. 8. There, although the groupings are disparate, so too are the average areal brightnesses of regions including

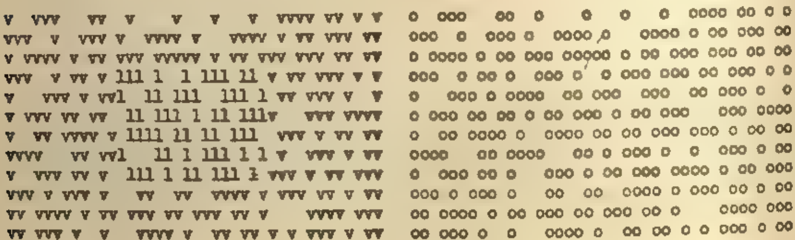

FIG. 7.

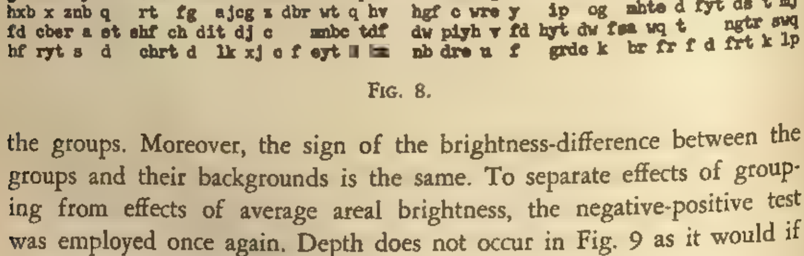

FIG. 8.

the groups. Moreover, the sign of the brightness-difference between the groups and their backgrounds is the same. To separate effects of grouping from effects of average areal brightness, the negative-positive test was employed once again. Depth does not occur in Fig. 9 as it would if mere groupings were capable of producing stereopsis.

The conclusion that grouping *per se* is not a sufficient basis for the occurrence of stereopsis must be accepted with some caution. It is possible, as it is for contour, that groupings can produce effective disparities but not in the presence of competing average brightnesses. Thus, grouping, contour, and brightness are all potential stimuli, but none is independently so. It is argued here that relative brightness-disparity can account for all

of the stereo-effects and, therefore, it is more parsimonious to consider it to be the basic stimulus-condition. This conclusion does not, however, follow inevitably from the evidence.

Demonstration V

In connection with stereograms such as those shown in Fig. 7 and 8, it may be argued that the depth-effect results from a fusion of disparate subjective contours surrounding clusters of letters. This notion suffers in the light of the last demonstration. The subjective-contour factor may also exist but, as intimated above, be unable to operate in the presence of brightness-induced rivalry across the subjective contours. There would, therefore, be a limit to the amount of tolerable rivalry beyond which

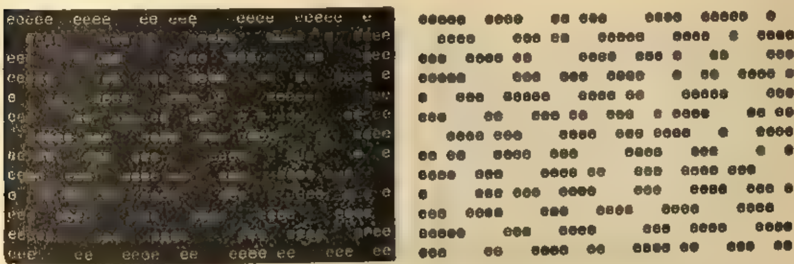

Fig. 9.

stereopsis would fail.[8] The fact that both letters and background in Fig. 9 were rivalrous may have caused the rivalry to exceed this limit in that stereogram. It is therefore of some importance to determine if the ostensive grouping effects are based upon subjective contours. One way to do this is to produce a pattern in which there are disparate groupings which are not bounded by disparate subjective contours.

The right half of the stereogram in Fig. 10 is a simple matrix of *l*s and *o*s distributed at random. The left half has a discriminable central region comprised of *v*s and =s. The *v*s of this region correspond to *o*s in the contralateral field, while the =s correspond to *l*s. The corresponding groups of *l*s and *o*s are shifted one letter-space toward one side, thus creating a grouping-disparity. The use of randomly placed *l*s and *o*s over the entirety of one side of this stereogram prevents the formation of a subjective contour surrounding both inner regions, while it does permit the presence of correlated groups of letters.

[8] J. E. Hochberg (Depth perception loss with local monocular suppression: A problem in the explanation of stereopsis, *Science*, 145, 1964, 1334-1335) provides evidence for this possibility.

We have had some difficulty in obtaining valid judgments of the direction of depth in this pattern when naïve *O*s viewed it in a mirror-stereoscope. As a matter of fact, only one of three trained *O*s was able to state the direction of depth correctly more often than he would by chance alone. It was found after exploratory studies that this situation could be alleviated for trained *O*s when one of two procedures was used. First, with a plus-1-Diopter lens before each eye, and with the patterns so mounted in the stereoscope that the optical path length was 1 m., the direction of depth was reported correctly 40 times out of 40 trials by all three trained *O*s. Secondly, and this is something of a mystery, we found that flickering one side of the stereogram in such a way that it was illuminated for 30 m.sec. and in darkness for 100 m.sec. also resulted in valid reports of

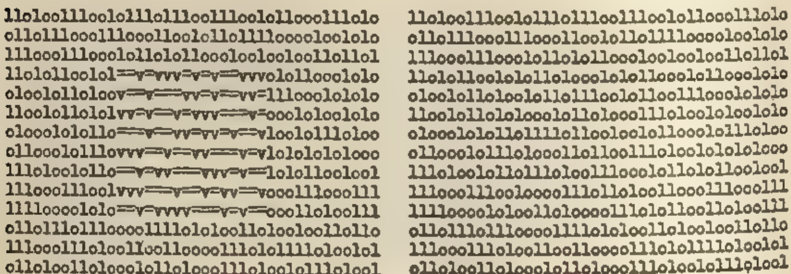

FIG. 10.

direction. This procedure was employed for an entirely different purpose, and the effect was noted accidentally. Further study showed that some naïve *O*s improved in their ability to discriminate the different effects of different crosses of disparity but, for the most part, only trained *O*s were able to employ the flicker in assessing depth.

The result with the lenses suggested that bringing accommodation and convergence into agreement might be necessary to produce stereopsis for inexperienced viewers of this pattern. To this end, the patterns were placed in register on the screen of a polarizing device and viewed through polaroid spectacles. Under this condition, six naïve *O*s had no trouble in detecting depth and correctly describing its direction. These *O*s were asked to range on the patterns in a manner described fully in a previous paper.[9] The average amount of depth thus obtained was 10.07 cm. at a viewing-distance of 150 cm. The magnitude and direction of this effect

[9] Kaufman, *op. cit.*, this JOURNAL, 78, 1965, 4ff.

would not be expected if inner regions of comparable size were simply rivalrous and non-disparate.

It is clear from this experiment that depth does not depend upon the presence of subjective contours. It may be due to the correlated groupings, or, alternatively, to the presence of disparity between regions which have correlated average areal brightnesses. Although =s and *v*s have the same brightness-values, they have different over-all contour-lengths in a unit-area of the display. This results in different average areal brightnesses. So, too, with the *l*s and *o*s. It should be possible, therefore, to vitiate the effect of a grouping disparity by interchanging =s and *v*s in such a way that =s correspond to *o*s and *v*s to *l*s, thereby changing the signs of the directions of areal-brightness change. That is, in fact, what happens.

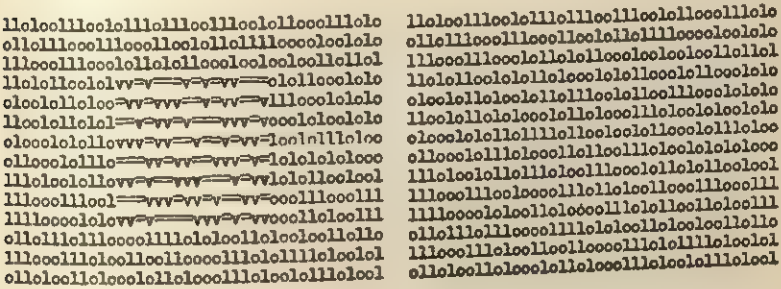

FIG. 11.

When the *v*s and =s in Fig. 10 are interchanged as shown in Fig. 11, the depth becomes ambiguous as verified by two trained *O*s, thus confirming the notion that grouping disparities are effective only insofar as they produce brightness-disparities of the sort described.

In another study, whose purpose is now obsolete, we placed *v*s and =s and *l*s and *o*s in two halves of a stereogram and surrounded them with randomly-selected letters of the alphabet. Both inner regions in this pattern were isolated statistically and could be discriminated monocularly. In this type of pattern, we found that the *v*s and =s could be interchanged and the depth-effect preserved. In view of the foregoing, our explanation is that the statistical isolation of the two inner regions created large areas whose average areal brightness was greater than that of their surrounds. Interchanging the letters did not affect this average, since the distribution of long-contour and short-contour letters was not affected significantly. Rearrangement, without systematic correlations between the two patterns, did destroy stereopsis, because it altered the distribution of

light and dark elements. We can see from this that the visual system is capable of reacting to some average of brightness over both large and small areas, and that disparity between these areas is the effective cue to stereopsis.

Conclusions

This paper represents an effort to trace the broad outlines of the stimulus-conditions relevant to stereopsis. Work is continuing and quantitative data are being gathered. The conclusions drawn here, then, are somewhat tentative, and they are intended only as suggestions for researchers in this newly reactivated field.

Our observations suggest that relative brightness-disparities are both necessary and sufficient stimuli for stereopsis. The brightness-levels of the stimuli need not be the same in the two eyes, nor do the contrast-ratios between the elements of a stereogram and their backgrounds. It is merely the signs of the brightness-differences between disparate elements and their backgrounds that must be the same. No doubt there are brightness-contrast differences which are so great as to offset this responsiveness to sign. These parameters are now under investigation.

The foregoing is consistent with the exciting results obtained by Treisman.[10] She found that as long as the sign of the brightness-change across a contour was the same in the two eyes, stereopsis could result. She also found that if the sign of a color-change was the same across the contour imaged in the two eyes stereopsis would result even if the sign of the brightness-change was not the same. This suggests that chromatic color may be a second dimension relevant to stereopsis. Treisman's stimuli did not, however, separate brightness-effects from contour-effects. It is very important to note that the patterns used here indicate that there need be no contour-similarity whatsoever between display-elements in order for them to be seen in depth as a result of brightness-disparity.

Assume that there is a phenomenally segregated patch in one eye which is, on the average, brighter than its background. If another phenomenally segregated patch of roughly the same size is presented disparately to the other eye, and if its average brightness is also greater than that of its background, then stereopsis may result. In fact, the segregation itself may be due to the average brightness-difference. Furthermore, the patch-size may range from that of a very small cluster of dots, as in a Julesz-pattern, to a large collection of typed letters of the alphabet. The size-limits within which effective brightness-disparities occur have yet to be defined, as has

[10] Treisman, *op. cit.*, 23-37.

the mechanism by which the visual system may be adjusted to the relevant size-dimension in the case of ambiguity. When the depth-information of a display is carried by small regions as well as by gross regions, the two consistent effects reinforce each other to make the depth-effect readily apparent. When the brightness-disparities in small local regions are in conflict with relative brightness-disparities in somewhat larger regions, then both can be responded to at different times.

We believe that Whittle's hypothesis about Helmholtz's negative-contrast patterns is substantiated.[11] That hypothesis, together with the postulate that brightness-disparity is the effective stimulus to stereopsis, makes it possible to deal with Helmholtz's demonstration and its failure in the Julesz negative-contrast pattern in terms of the same theoretical model.

The notion that contour is essentially irrelevant as a basis for stereopsis is of some importance. It has been suggested that, even if rivalry is present in a stereogram, the contours can still fuse to produce stereopsis.[12] It should be evident from these demonstrations that not only is rivalry present but that there are also no contours, subjective or real, which could conceivably fuse in some projection-field. Moreover, the rivalry can be seen to occur between the different surface-colors employed in the patterns. Sometimes one sees a black letter and then a gray letter of the same shape in the same place. We cannot, therefore, accept the notion that rivalry is complementary to stereopsis in the sense that its occurrence between the relevant disparate dimensions destroys stereopsis.[13] At the very least, rivalry is a relatively independent phenomenon. It is, however, more likely that rivalry reflects the nature of the processing performed by the nervous system to ascertain the depth-relations.

---

[11] Whittle, *op. cit.*, 50-51.

[12] Treisman, *op. cit.*, 29-30.

[13] Gerhard Ronne, The physiological basis of sensory fusion, *Acta Ophthalmologica*, 34, 1956, 1-26.

THE JUDGMENT OF SIZE

By Martha Teghtsoonian, Smith College

When a 1000 ~ tone is increased tenfold in intensity (sound-energy), the average listener will report that the loudness of the tone has approximately doubled, but an acoustics engineer who is thoroughly familiar with sound-level meter-readings may correctly report the ratio of sound-energies as 10:1. Although this distinction between judgments of loudness and judgments of physical intensity has been recognized for many years, the exact form of the function relating these two measures has only recently received adequate specification.[1]

A rather different treatment of the problem of the judgment of size has developed. Although various writers have suggested the value of a distinction between 'apparent size' and 'physical size,'[2] many investigators have been primarily concerned with the capacity of *Os* to make accurate judgments of physical size despite changes in critical aspects of the viewing situation. In particular, interest has focused on the role of distance in judgments of size. Hence, a central issue has been the phenomenon of size-constancy. Because judgments of size may be veridical in size-constancy experiments, the assumption is sometimes made that physical and apparent size stand in a simple 1:1 relation, and there is some direct evidence for this view.[3] Yet Ekman has reported results showing that judged size of circles is related to their physical size by a power function with an exponent of 0.86, or, in other words, that judged size increases somewhat more slowly than the stimulus.[4] Williams has reported an exponent

* Received for publication November 15, 1963. This paper is based, in part, on a doctoral dissertation submitted to Radcliffe College. Experiment I was carried out at Harvard University, Experiments IIa and III at the Eastern Pennsylvania Psychiatric Institute and Experiment IIb at the Ontario Hospital, St. Thomas, Ontario, Canada.

[1] S. S. Stevens, The direct estimation of sensory magnitudes—Loudness, this JOURNAL, 69, 1956, 1-25.

[2] R. B. Joynson, The problem of size and distance, *Quart. J. exp. Psychol.*, 1, 1949, 119-135; R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, 1956, 485.

[3] S. S. Stevens and E. H. Galanter, Ratio scales and category scales for a dozen perceptual continua, *J. exp. Psychol.*, 54, 1957, 377-411; Gösta Ekman and K. Junge, Psychophysical relations in visual perception of length, area, and volume, *Scand. J. Psychol.*, 2, 1961, 1-10.

[4] Ekman, Two generalized ratio scaling methods, *J. Psychol.*, 45, 1958, 287-295.

of 0.81 for a variety of two-dimensional figures.[5] Stevens and Guirao obtained an exponent of 1.0 for straight lines but 0.7 for squares.[6]

The first experiment to be reported tests the possibility that the relation between judged and physical size depends on the task which is given to *O*. He may in one case be asked to estimate how large a given stimulus *is;* alternatively, he may be asked to judge how large the stimulus *looks.* A difference between the results under the two sets of instructions would constitute grounds for defining a construct of apparent size analogous to that of loudness. The second experiment deals with the role of judgments of length in the judgment of size. The third experiment deals with the number of spatial dimensions of a stimulus as a determiner of its apparent size.

Experiment I

Judgments of size were obtained by the method of magnitude-estimation for each of two separate sets of instructions. One set required *O* to estimate the physical areas of two-dimensional stimuli; the other required *O* to judge how large these stimuli appeared.

Stimuli. The stimuli were circles cut from heavy white cardboard. To obtain a judgment, a given figure was suspended centrally on a large black-cloth screen, 4 × 5 ft., which was approximately 10 ft. from the seated *O*. Nine circular figures were used, ranging from 1-9 in. in radius in 1-in. steps; the range of areas was from 3.14-254 sq. in. The standard stimulus, the middle figure in the series, was 78.5 sq. in.

Instructions. Two instructions were used. Since the difference between them is central to the argument of this study, both are given in their entirety.

Objective-area. I am going to show you some circles of different areas, and I want you to tell me how large they are, in terms of area. First I shall show you a circle whose area will arbitrarily be called 10; then I'll show you others; to each one assign a number which is proportional to its area in terms of square units. For instance, a circle whose area is twice as great as that of the one called 10 would be called 20; one whose area is half as great would be called 5; and so on. Don't worry too much about being consistent or trying to remember what you assigned a circle before; just judge each one as it comes along. Obviously there are as many numbers above 10 as you could want to use, but there are also as many below 10, since you can use fractions, 1/2 or 1/7, for instance. Any question? Then we'll start. [Standard displayed.] This is the circle whose area is called 10. [Standard removed, first stimulus displayed.] If the first circle was 10, what would you call this one?

Apparent-size. I'm going to show you some circles of different sizes, and I want you to tell me how big they look to you. First, I shall show you a circle whose size will be called 10, arbitrarily; then I'll show you others; to each one assign a number which

[5] R. L. Williams, *Statistical Symbols for Maps: Their Design and Relative Values,* 1956, 64-74.

[6] S. S. Stevens and Miguelina Guirao, Subjective scaling of length and area and the matching of length to loudness and brightness, *J. exp. Psychol.,* 66, 1963, 177-186.

is proportional to its size as it appears to you. For instance, a circle which looks twice as big as the one called 10 would be called 20, one which looks half as big would be called 5, and so on. [These instructions were completed as in the instructions for Objective-Area.]

Observers. The 10 *O*s were graduate students in psychology.

Procedure. Each *O* was tested twice, once with each set of instructions. For half the *O*s, one set of instructions was used first; for the remaining half, the other set was used first. The shortest interval between sessions was one day.

The method employed was the method of magnitude-estimation, which has been described in detail by Stevens.[7] Two judgments were obtained for each figure in the series in each session. The order of presentation was random, with the restriction that the same figure could not immediately succeed itself and that the standard could not be the first presented. The standard was presented and identified at the beginning of each session; it was twice presented subsequently for judgment but not identified.

Results. A correction was applied to all judgments of each *O*, when required, to make the arithmetic mean of his judgments of the standard-stimulus equal to 10. Geometric means were computed for each of the 9 stimuli (based on 20 scores per stimulus). The two functions thus obtained are shown in Fig. 1, whose coördinates are logarithmic. The straight lines were fitted by the method of least-squares. Two features of these results should be especially noted. First, both sets of instructions lead to judgments which are reasonably described as power functions of physical area. Secondly, while judgments of area follow a power function with an exponent of 1.03 ($\sigma_{yx} = 0.06$), the exponent of the power function describing size-judgments is 0.76 ($\sigma_{yx} = 0.05$).

As a further test of the hypothesis that instructions influence the value of the exponent, separate functions were determined for each *O*, and in each case the exponent was computed by the method of least-squares. In 9 out of the 10 cases, the exponent was greater under objective-area instructions than under apparent-size instructions.

Discussion. The results support the view that a valid distinction may be drawn between judgments of physical area and judgments of apparent size. While it is evident that *O*s may make accurate area-estimates, the data indicate that apparent size grows at a slower rate than physical area: a circle twice the area of the standard circle looks, on the average, about 1.7 times as large; a circle 10 times the area looks only about 6 times as large. Of particular interest is the finding that instructions play an important role in judgment. Any study of judged size which employs

---

[7] Stevens, *op. cit.*, 5-6.

instructions that are ambiguous in defining *O*'s task is likely to yield ambiguous results.

In one respect, however, the results are paradoxical. If *O* can judge physical area accurately, why are his judgments of size not linearly related to physical area? The problem is perhaps more usefully stated by turning the question around: if size appears to *O* to grow more slowly

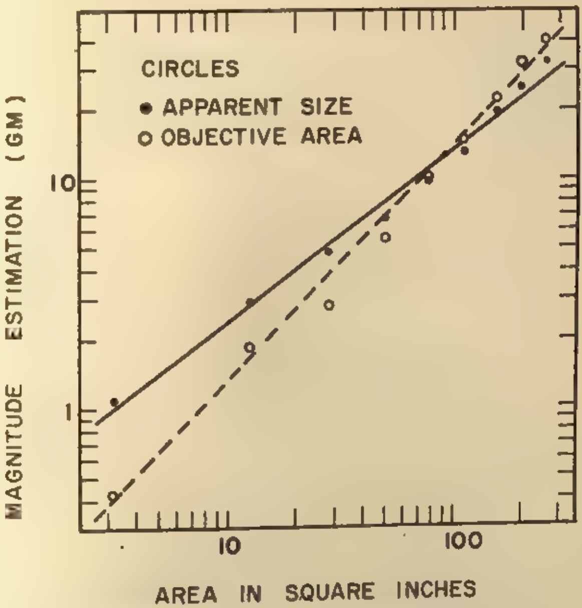

FIG. 1. MAGNITUDE-ESTIMATIONS FOR CIRCLES UNDER OBJECTIVE-AREA AND APPARENT-SIZE INSTRUCTIONS

Every point is the geometric mean of 20 judgments, two from each *O*. The coördinates are logarithmic. The straight lines represent least-squares solutions: the slope is 0.76 for apparent size (closed circles), 1.03 for objective area (open circles).

than physical area, how can he nevertheless be capable of accurate judgments of physical area? *O*s were asked, following Experiment I, how they had made physical-area judgments. Most replies indicated that the judgments were based on estimating a linear dimension of the stimulus figure and squaring it. These data suggest that *O*'s ability to make accurate area-judgments depends on (1) his ability accurately to estimate length, and (2) his knowledge that the area of a two-dimensional figure is proportional to the square of a linear dimension. Thus, a knowledgeable *O*

who correctly judges one circle to be twice the diameter of another *knows* its area to be four times as great, even though it does not *look* that large. He is like an *O* who inspects an example of the vertical-horizontal illusion and then measures the two axes with a ruler; he can tell you that they are physically equal even though they look unequal.

EXPERIMENT II

In Experiments II, the two factors hypothesized to account for accurate physical-area estimates were subjected to analysis. That *O*s can estimate length in inches quite accurately is known.[8] More important to the present argument is the fact that apparent length will also be related in a linear fashion to physical length. If, under physical-length instructions, an exponent of 1.0 is obtained, but, under apparent-length instructions, the exponent is less (or greater) than 1.0, the paradox of accurate area-judgments is not resolved; rather, a second paradox is added to the first. Accordingly, if the hypothesis is correct, physical-length and apparent-length instructions will both result in exponents of 1.0.

The second factor, *O*'s knowledge of the relation between linear dimension and area, was approached by eliminating the possibility of judgments based on this knowledge. The stimuli consisted of a series of irregular polygons which varied, not only in area, but in shape. Since there is no regular relation within such a series of non-similar figures between linear dimension and area, *O* should be able to base his judgments only on the apparent size of the stimuli. The prediction was made, therefore, that instructions for objective area and apparent size would result in the same kinds of judgment, with exponents like those of regular two-dimensional figures.

Experiment II a: Straight lines. A group of 10 *O*s, technical and professional personnel in a clinical research unit, were tested as in Experiment I, except that the stimuli were nine straight lines, ranging in length from 0.5-33 in.; the standard line was 3.5 in. Slides were used to project a white line, horizontally oriented, against the same black background as was previously used.

Results. As Fig. 2 shows, there is no difference in the slopes of the functions obtained under the two sets of instructions. The exponent for apparent-length instructions is 0.98 ($\sigma_{yx} = 0.02$), and for physical-length instructions it is 1.02 ($\sigma_{yx} = 0.03$). The straight line fitted by eye in Fig. 2 has a slope of 1.00. The individual exponents have a tendency to be slightly higher for physical-length instructions although with two ties,

---

[8] Stevens and Galanter, *op. cit.*, 378.

the difference is not significant ($P > 0.05$) by a sign-test. The results support the conclusion that *O*s make veridical judgments of a linear dimension regardless of the instructional set which is provided.

Experiment II b: Irregular polygons. Nine irregular polygons were constructed by Method 2 (for closed-contour, angular shapes) described by Attneave and Arnoult.[9] Two random orderings (Series *A* and *B*) of these nine stimuli were used, and for each ordering the nine forms were projected into physical areas with approximately

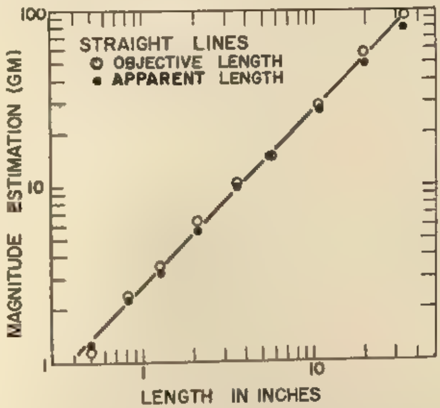

Fig. 2. Magnitude-Estimations for Straight Lines under Objective-Length and Apparent-Length Instructions

Every point is the geometric mean of 20 judgments, two from each *O*. The coördinates are logarithmic. Closed circles represent apparent length and the open circles represent objective length. The straight line, fitted by eye, has a slope of 1.00.

equal logarithmic spacing over a two-cycle range. The areas, in square inches, ranged from 2.0-200; a figure of 20 sq. in., the middle stimulus in each series, served as the standard. The *O*s were 10 graduate students in psychology, 5 for each series, and the procedure was identical in all respects to that employed in Experiment I.

Results. Geometric means were obtained for the 5 *O*s judging Series *A* and for the 5 *O*s judging Series *B*. The particular ordering of polygons had no differential effect on judgments of size or area. For apparent size, the exponent for Series *A* was 0.80, and, for Series *B*, 0.81, σ_{yx}-values of 0.04 and 0.05, respectively; for estimated area, Series *A* was 0.78 and

[9] Fred Attneave and M. D. Arnoult, The quantitative study of shape and pattern perception, *Psychol. Bull.*, 53, 1956, 452-471.

Series *B*, 0.76, with σ_{yx}-values of 0.04 and 0.05, respectively. The data for both series were consequently combined.

These combined data show no effect of instructions on the judgment of irregular, non-similar polygons. The exponent under instructions for apparent-size is 0.83 ($\sigma_{yx} = 0.04$); for objective area instructions, 0.80 ($\sigma_{yx} = 0.04$). The geometric means are shown in Fig. 3; the line, fitted by

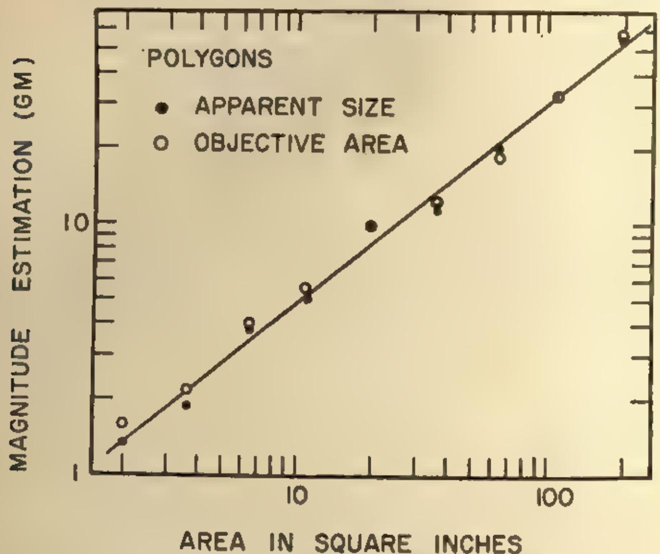

FIG. 3. MAGNITUDE-ESTIMATIONS FOR IRREGULAR, NON-SIMILAR POLYGONS UNDER OBJECTIVE-AREA AND APPARENT-SIZE INSTRUCTIONS

Every point is the geometric mean of 20 judgments, two from each *O*. The coördinates are logarithmic. Closed circles represent apparent size and open circles represent objective area. The straight line, fitted by eye, has a slope of 0.81.

eye, has a slope of 0.81. Of the 10 individual exponents, the exponent for objective area was greater than that for apparent size in three cases, and there was one tie; the difference was not significant by a sign-test ($P >$ 0.05).

Discussion. The results support the view that veridical areal judgments depend on the availability of a linear dimension common to all of the figures in the series. Deprived of such a dimension, *O*s behave no differently under objective-area instructions than under apparent-size instructions. This result stands in marked contrast to the results obtained for circles.

EXPERIMENT III

The exponents obtained for the two two-dimensional forms are close to 0.8, while the exponent for straight lines is about 1.0. A number of other two-dimensional figures have been presented to *O*s under apparent-size instructions, and the results suggest that form has a relatively small influence on the exponent of the power function. (The exponents for squares, parallelograms, equilateral triangles, irregular similar polygons, circles, and various ellipses with major-to-minor axis-ratios of 6:5, 6:3, and 6:1 range from 0.60-0.90, with most falling in the interval 0.76-0.81.)

The possibility should be considered, however, that the dimensionality of the stimulus is a critical factor in determining the exponent of the power function for apparent size, and to test that possibility some work was done with cubes and octahedrons.

Stimuli. Two series of cubes were constructed of heavy white cardboard: the first consisted of 10 cubes covering a range of volumes from 1.00-1000 cu.in. with a standard cube of 76.8 cu.in.; the second series consisted of eight cubes, covering a range of 1.00-145 cu.in., with a standard cube of 11.4 cu.in.

Two series of octahedrons, also constructed of white cardboard, were used: the first, with 10 stimuli, ranged in volume from 0.92-974 cu.in. with a standard of 42.8 cu.in.; the second, with 7 stimuli, ranged from 2.52-198 cu.in. with a standard of 20.2 cu.in.

Procedure. The procedure and apparent-size instructions of Experiment I were employed, with one exception: instead of the figures being displayed in front of a black screen at a given distance from *O*, they were placed on a desk in front of him and he was allowed to handle and rotate them if he wished.

Observers. Ten *O*s were asked to make apparent-volume judgments of each set of cubes and set of octahedrons (40 *O*s in all). The *O*s who judged cubes were graduate students in psychology while those who judged octahedrons were technical and professional personnel in a clinical research unit.

Results. The data for the four groups are shown in Fig. 4, the placement of the functions on the *y*-axis having been chosen to bring all four within the same numerical range. For the 3-cycle series of cubes, the exponent is 0.67 ($\sigma_{yx} = 0.02$); for the 2 cycle series, 0.72 ($\sigma_{yx} = 0.02$). For the 3-cycle series of octahedrons, the exponent is 0.65 ($\sigma_{yx} = 0.04$); for the 2-cycle, it is 0.74 ($\sigma_{yx} = 0.04$). The straight line fitted to the points in Fig. 4 has a slope of 0.70.

Discussion. These results provide at least tentative support for the view that judged size is a power function of physical size, with the exponent a decreasing function of the number of spatial dimensions of the stimulus. It is possible, of course, that the lower exponents for two- and

three-dimensional figures are an artifact of averaging data from two kinds of *O*s: those who make veridical area- and volume-judgments, and those who make veridical judgments of a linear dimension of two- and three-dimensional figures. If that were the case, the distribution of individual exponents for two-dimensional figures should peak at 0.5 and at 1.0; for the three-dimensional figures, at 0.3 and 1.0, but Fig. 5 shows that the distributions are unimodal. There are some exponents at 0.5 for two dimensional figures, and some at 0.4 for three-dimensional figures, and there

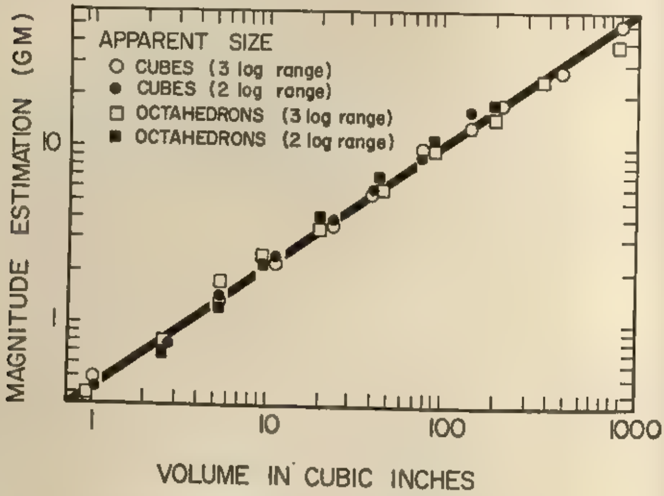

FIG. 4. MAGNITUDE-ESTIMATIONS FOR CUBES AND OCTAHEDRONS UNDER APPARENT-SIZE INSTRUCTIONS

Every point is the geometric mean of 20 judgments, two from each *O*. The coördinates are logarithmic. Circles represent cubes, and squares represent octahedrons. Open symbols represent 3-long range series and closed symbols represent 2-log range series. The numerical values on the *y*-axis are arbitrary: the four functions have been superimposed for easier comparison. The straight line fitted by eye, has a slope of 0.70.

are exponents of 1.0 for both; but the modal value for straight lines is 1.0, for two-dimensional figures is 0.8, and for three-dimensional figures is 0.6.[10] Thus, the exponents derived from grouped data are a fair reflection of the distributions of individual exponents. Pending investigation of additional three-dimensional figures, then, it seems likely that for

[10] These include all individual data. Some *O*s served more than once. If each *O* is allowed to contribute to the distribution only once, the picture is not markedly changed: the main difference is that the distribution for two-dimensional figures develops a second mode at 0.9.

any given number of dimensions, the form of the figure is of relatively little importance to the value of the exponent. (Fig. 5 includes the data reported here and additional data from a series of studies in which *O*s made apparent-size judgments of lines and a variety of two-dimensional figures.)

The present results are at variance with those reported by Ekman and Junge, who studied ratio-estimations of straight lines, squares, and cubes and found exponents of about 1.0 for all figures.[11] Their method differed

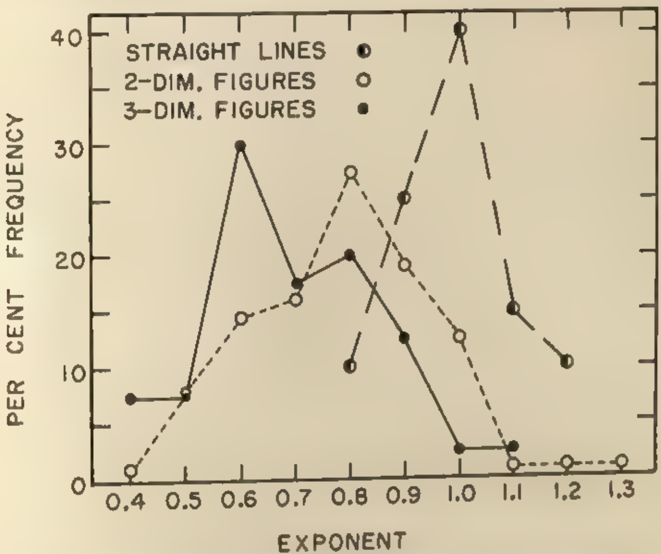

FIG. 5. FREQUENCY-DISTRIBUTIONS OF INDIVIDUAL EXPONENTS FOR STRAIGHT LINES AND FOR TWO- AND THREE-DIMENSIONAL FIGURES
The frequencies are converted to percentages: straight lines (half-open circles) are based on 20 *O*s; two-dimensional figures (open circles) are based on 130 *O*s; three-dimensional figures (filled circles) are based on 40 *O*s. The exponents are least-square fits to individual data from magnitude-estimations under apparent-size instructions.

from the one used here, and they did not specify the exact form of their instructions. Since their *O*s were men—school teachers, who, it may be assumed, were knowledgeable about the relationship between linear dimension and area or volume—it would not be surprising, in view of the results of Experiment I, if, in the absence of specific instructions, the *O*s attempted to make their ratio-estimates correspond to estimated area and volume.

[11] Ekman and Junge, *op. cit.*, 2-40.

It is noteworthy that in 1932, Croxton and Stein reported a study in which *O*s were asked to estimate the sizes of smaller figures in relation to larger figures.[13] Bars, circles, squares, and two-dimensional projections of cubes were employed. Although only 'error' measures were reported, it is interesting that no difference was found between circles and squares; that both were more subject to error than were bars; and that cubes were most subject to error. This outcome is compatible with the present results.

Summary

Judged size, scaled by the method of magnitude-estimation, is related to physical size by a power function. The exponent of the power function may take a value of 1.0—that is, judged size may be in direct correspondence to physical size—only under special circumstances: (1) for judgments of linear dimensions, and (2) when area-ratios can be successfully determined on the basis of judgments of length *and* physical area is estimated. When *O* is asked to judge the apparent size of two- and three-dimensional figures, the exponent of the power function is appreciably less than unity. For two-dimensional figures, this value is about 0.8, and for three-dimensional figures, it is about 0.7. Neither value appears to be greatly influenced by the form of the figures studied.

[13] F. E. Croxton and Harold Stein. Graphic comparisons by bars, squares, circles and cubes, *J. Amer. stat. Assn.*, 27, 1932, 54-60.

ORGANIZATIONAL FACTORS IN LEARNING AND REMEMBERING: FUNCTIONAL UNITY OF THE INTERPOLATED TASK AS A FACTOR IN RETROACTIVE INTERFERENCE

By Sandor B. Brent, Wayne State University

This study is concerned with the organizational level of the interpolated learning (*IL*) material as a factor in retroactive interference (*RI*). Empirical studies and theoretical formulations of *RI* have tended to focus upon four variables: *temporal intervals* between original learning (*OL*), interpolated learning (*IL*), and relearning (*RL*); *number of items* in the *OL* and *IL*; *number of trials* on the *OL* and *IL*; and *similarity of content* in the *OL*- and *IL*-materials.[1] It seems reasonable to assume, however, that the degree to which the elements of *IL* are "free" to interfere with recall of the *OL*-material should depend on the *internal structure* of the *IL*-task as well. The present study was designed to investigate the relationship between the degree of *functional unity* of the *IL*-material and the magnitude of *RI*, when the content of both learning tasks, their temporal relationships, and the number of trials on each are held constant.

The Concept of Functional Unity

"Functional unity" here refers to the degree to which the functions of each of the elements in a serial word-list contribute to and are determined by some supraordinate functions (or functional rules) governing *S*'s organization of the list as a whole. This concept of Functional Unity is derived from the familiar emphasis of gestalt psychology upon the interdependence of parts and whole,[2] with the addition of the organismic-developmental assumption that the degree of interdependence may vary depending upon the type of organizational structure which is established by *S*.[3]

* Received for publication January 18, 1965. This research was supported, in part, by NSF Grant GB1579, and in part with funds made available by the Carnegie Foundation, through the Wayne State University Center for the Study of Cognitive Processes. The author expresses, his appreciation to Quine Brengle, III, Elaine Feldman, and Deborah Zimmerman for their assistance in this experiment.

[1] See, for example, Leo Postman and Donald Riley, Degree of learning and interserial interference in retention, *University of California Publication in Psychol.*, 8, 1957 (No. 4), 271-396.

[2] See Kurt Koffka, *Principles of Gestalt Psychology*, 1935, 174-176, 568-570.

[3] See Heinz Werner and Bernard Kaplan, *Symbol Formation*, 1963, 3-25.

While different levels of functional unity (*FU*) may be achieved by manipulating any one or more of a variety of different modes of organization imposed by *S* upon a stimulus-list, the present study was specifically concerned with the effects of a "sentential" mode of organization upon lists of meaningful words.

Within this context, a list with an extremely *high* degree of *FU* is exemplified by the following series of words: *he went to the movies yesterday*. If one recognizes this sequence as a complete and meaningful English sentence, and grasps the meaning of the sentence, the identity and location of each of the elements in the list may be subsequently reconstructed with a high degree of accuracy. If, for example, after having learned this list one were confronted with the sequence *he went to movies the yesterday,* one would know almost immediately that the elements in the fourth and fifth serial positions were reversed. Likewise, confronted with the sequence *he went to the ? yesterday,* one would know with a high degree of certainty that the word "movies" was the missing element. In summary, in such a list, "conveying the meaning" of the sentence may be thought of as a supra-ordinate function served by the list as a whole. The "meaning conveyed," on the other hand, may be thought of as a "rule" for identifying and ordering the elements within the list. This rule may be seen to be both determined by, and itself a determinant of, the specific functions of the elements which must occupy each of the serial positions within this list.

A list with a relatively *low* degree of *FU* is exemplified by the following, somewhat *different ordering,* of exactly the same words: *the yesterday he to went movies*. If, after having learned this list, one were confronted with the sequence *the yesterday to he went movies,* one would have no way of determining with any high degree of certainty that the third and fourth elements had, in fact, been interchanged. Likewise, confronted with the sequence *the yesterday ? to went movies,* one could not with any degree of certainty reconstruct the missing element.

Clearly, the difference between the high- and the low-*FU* conditions cited above would increase with increasing list-length.

In terms of experimental manipulation, *FU* may be hypothesized to derive from any characteristic of the stimulus-list which helps *S* (1) to *differentiate systematically* the entire list into *functionally integral subunits*, and (2) to *relate* these subunits to each other in some regular fashion.

It follows from the foregoing discussion that a high-*FU* list should be less *subject to* various types of memory-interference than a comparable low-*FU* list. Since the function(s) of the element which occupies each serial position in a high-*FU* list is highly differentiated from the function of the element which occupies every other serial position in the list, such a list would be expected to be highly resistant to intra-serial interference. Since the function of each element is also specifically related to the functions of all of the other elements within the list by a supra-ordinate rule governing the list as a whole, such a list should also be highly resistant to inter-serial interference. The well-known fact that organized material is both easier to learn and easier to remember than relatively "disorganized" material of comparable length seems to add some empirical support to this hypothesis.[4]

[4] See, for example, R. S. Woodworth, *Experimental Psychology*, 1938, 60-61; William Epstein, The influence of syntactical structure on learning, this JOURNAL, 74, 1961, 80-85.

The question which remains to be answered is whether high-*FU* material is also less likely to be a source of interference into adjacently learned material. On purely theoretical grounds, it seemed reasonable to assume that the higher the degree of integration of the elements within a given list, the less "available" they would be for interference into adjacently learned material. The following experiments were conducted to investigate this question within a classical *RI* paradigm.

Experiment IA

The first experiment investigated the hypothesis that an *IL* list of *high-FU* would result in less *RI* than an *IL* list of *low-FU*. If *high-FU* did make the items within a list less available as a source of interference into adjacently learned material, then this hypothesis should be supported.

Method. A classical *RI*-paradigm with serial anticipation was used. All *S*s learned a standard *OL*-list to one perfect anticipation. *S*s were then alternately assigned either to the *high-FU* or *low-FU IL*-condition. After eight anticipatory trials on their respective *IL*-list, all *S*s relearned the *OL* to criterion.

Materials. The *OL*-list was composed of 18 common English words arranged in a "meaningless" order: *fellow it ran last expected race fifty this is to yesterday tomorrow the again he do yards in.* The order of the words was the same for all *S*s.

In the *low-FU* condition, these same 18 words were rearranged in yet another "meaningless"order: *tomorrow last yesterday yards it again he fellow to do is expected race fifty in this the ran.*

In the *high-FU* condition, the words were rearranged to form two meaningful English sentences: *this fellow ran fifty yards in the last race yesterday tomorrow he is expected to do it again.*

Subjects. There were 14 *S*s in each of the two experimental groups. The *S*s were students in an Introductory Pyschology course.

Procedures. Lists were presented on an *MTA-100* teaching machine at the rate of 4 sec. per item. Each list began with an asterisk and ended with a blank interval.

Instructions. At the beginning of the experiment, the *S*s were instructed as follows:

As I indicated [on the appointment card], this is a verbal learning experiment. We'll be doing several things. The first thing is to learn a series of words. The words will appear one at a time in that slot in the memory-drum. Your task is to memorize the words in the order they appear. I will know you have done this when you can anticipate what the next word is going to be every time. For example, if the list contained three words, *a, b,* and *c,* after you had seen them once, when *a* appeared you would say *b,* and when *b* appeared you would say *c,* and so forth. Do you understand? You will go over the list until you can anticipate all of the words correctly in a single time through the list.

Another thing—if you think you know a word, by all means guess—you might be right. Do not get discouraged; it is a difficult task. You will miss a lot at first and then all of a sudden it will get much easier. Do you have any questions?

Now, the first time just look at the words and try to remember them. Then, when the asterisk appears, say the first word if you remember it and so forth.

After completion of the *OL*-task, the *S*s were instructed as follows:

Fine. Now I'm going to give you a list of the *same* words but they will be in a different order this time. After you have seen the list once, begin to anticipate. Do you understand?

After completion of the *IL*-task, they were instructed as follows:

Good. Now I want you to relearn the original list of words. This time I want you to anticipate right from the beginning. If you have any idea be sure to guess. Do you remember the first word?

Results. The two experimental groups did not differ significantly on trials to criterion during *OL*.

RI was indexed by two different measures: mean errors on first *RL* trial, and total *RL* trials to criterion. Since these two scores were highly correlated and yielded essentially the same results, only errors on first *RL* trial are reported here.

Errors on first *RL* trial average 15.5% for the *high-FU* group and 57.2% for the *low-FU* group. A one-tailed Mann-Whitney U-test was significant beyond the 5% level ($U = 3$, $n_1 = n_2 = 14$).

These results showed (1) that level of *FU* did play a significant role in the degree to which the two *IL* lists interfered with recall of *OL*, and (2) that the *high-FU IL* list resulted in significantly less *RI* than the *low-FU* list, as hypothesized.

Experiment IB

The second experiment was directed toward determining whether the inverse relationship between *FU* and *RI*, observed in Experiment IA, was *continuous* over *intermediate* degrees of *FU*. If the relationship is continuous, an *IL*-list representing an *intermediate* degree of *FU* should result in an intermediate amount of *RI*, *i.e.* intermediate between those obtained under the two extreme *FU*-conditions investigated in Experiment IA.

Method. The method and procedures were identical with those of Experiment IA except that in this case only one group of *S*s was needed, since the results of the previous experiment provided the basis for comparison.

Materials. The *IL*-list consisted of a third rearrangement of the 18 words used in the *OL*-list of Experiment IA. Here, however, an attempt was made to increase the likelihood that subportions of the list would be perceived as meaningful *phrases* while the list as a whole would *not* readily be perceived as a complete sentence. The following order was used for the *IL*-list of *intermediate-FU: fifty yards this fellow yesterday in the last race ran is tomorrow he expected it again to do.*

Results. This group did not differ significantly from either of the previous two groups on *OL*-trials to criterion.

A *U*-test showed that the 50% *RI* obtained under this intermediate condition did not differ significantly from that obtained under the *low-FU* condition of Experiment IA ($U = 76$, $n_1 = n_2 = 14$), as would be expected if the inverse relationship between *RI* and *FU* were continuous. This *intermediate* condition did, however, result in significantly more *RI* than did the *high-FU* condition of Experiment IA ($U = 10.0$, $n_1 = n_2 = 14$).

Discussion. The results of Experiment IB did not support the hypothesis of a continuous inverse relationship between *FU* and *RI*. On the contrary, the results seemed to indicate a *discontinuous* relationship, such that the *highest* level of *FU* (complete sentences) differs from both the *intermediate* level (phrases) and the *low* level (randomized words), while the latter two do not differ significantly from each other.

The lack of difference between these two levels of *FU* might have stemmed from the fact that *S*s (for some unknown reason) did not "perceive" the difference between the *intermediate-* and *low-Fu* condition which *E* attempted to build into the two lists. If this hypothesis were valid, there should also have been no difference in the rate at which the two lists were learned during *IL*. Two measures of rate of learning were computed: percentage of anticipations correct on *last* (eighth) *IL* trial, and percentage of anticipations correct on *all IL* trials. These two measures were highly correlated and lead to essentially the same conclusions, viz. that the *intermediate-IL* list was learned significantly faster—was significantly easier to learn—than the *low-IL* list. (For mean percentage of correct anticipations on last *IL* trial, $U = 59.5$, $n_1 = n_2 = 14$). This fact shows that during *IL* the *S*s *were* able to utilize the difference between the *intermediate-* and the *low-FU* list and hence refutes the hypothesis that *S*s did not "perceive" this difference. The superiority of learning the *intermediate-FU* list also supports the assumption that this list did, in fact, represent a higher degree of *FU* than did the *low-IL* list.

Table I summarizes the findings of Experiments IA and IB together. Both *rate of IL* (in terms of mean percentage correct on last *IL*) and *amount of RI* (in terms of mean percentage of errors on first *RL* trial) are shown for each level of *FU*. The bracket ties together those conditions which *did not* differ significantly. These results indicate that variations in the *internal* structure of the *IL* can, at one and the same time, (a) have a marked effect upon the rate at which the *IL* items are acquired, and (b) have no significant effect upon the degree to which the learning of the *IL* structure *interferes with recall* of the *OL* structure.

Because the discrepancy between *rate of IL* and *amount of RI* was so strikingly different from any expected on the basis of existing experimental of theoretical literature on retroaction, and because the crucial evidence in support of this finding was based upon a *post hoc* analysis of the *IL* learning data, a second experiment was conducted.

Experiment II

The second experiment was conducted to determine the reliability of the findings of Experiment I. In the present experiment, however, an attempt was made to construct an *intermediate-FU* list which was *as similar as possible* to the *high-FU* list and *as different as possible* from the *low-FU* list. It was hoped that this would increase the likelihood of finding a difference in the amount of *RI* between the *low* and *intermediate* conditions, hence reduce the likelihood of replicating the *RI* findings of Experiments IA and IB.

TABLE I

Results of Experiments IA and IB

Level of *FU*	*N*	Rate of *IL*: Mean percentage of correct anticipations on last *IL* trial	Amount of *RI*: Mean percentage of errors on first *RL* trial
High	14	100.0	15.5
Intermediate	14	91.2	50.0 ⎤
			n.s. ⎥
Low	14	82.1	57.2 ⎦

Note.—Bracket connects conditions which *did not* differ significantly; a one-tailed Mann-Whitney U-test, 5% level of confidence is used throughout.

Method. The *OL*-list and all other procedures were identical with those of Experiment I, with the exception of the principle used to construct the three levels of *IL*-material.

Both the *high-* and the *intermediate-FU* lists in this experiment were variants of the *high-FU* list of Experiment IA—*i.e.* the 18 words arranged in two complete sentences. To this basic list, four new words—*young, five, up, hill*—were added.

For the *high-FU* condition, these words were added in positions which allowed them to be readily integrated into the meaning of the list as a whole: *this young fellow ran fifty five yards up hill in the last race yesterday tomorrow he is expected to do it again.*

For the *intermediate-FU* condition, these four words were added in positions in which they could *not* be readily integrated into the meaning of the list as a whole: *this fellow ran fifty up yards in the hill last race yesterday tomorrow young he is expected to do five it again.*

Finally, for the *low-FU* condition, the same four words were added at random to

the *low-FU* list used in Experiment IA; *tomorrow last yesterday yards up it again he hill fellow to do is young expected race fifty in this five the ran.* Thus, each *IL*-list now contained 22 words, while the *OL*-list still contained 18 words.

In the present experiment, there were 12 different *Ss* in each of the three experimental groups.

Results. Again there was no significant difference between groups on *OL* trials to criterion.

Table II shows the *rate of IL* (in mean percentage of correct anticipations on last trial of *IL*) and *amount of RI* (in mean percentage of errors on first *RL* trial) for each level of *FU*. The *intermediate-FU* list was learned significantly more rapidly than the *low-FU* list ($U = 24$, $n_1 = n_2 = 12$). The *intermediate-FU* list did not differ, however, from the *low-FU* list in the amount of *RI* generated ($U = 70$, $n_1 = n_2 = 12$).

TABLE II

RESULTS OF EXPERIMENT II

Level of *FU*	*N*	Rate of *IL*: Mean percentage of correct anticipations on last *IL* trial	Amount of *RI*: Mean percentage of errors on first *RL* trial
High	12	100.0 ⎤ n.s.	19.4
Intermediate	12	98.1 ⎦	56.1 ⎤ n.s.
Low	12	80.0	53.3 ⎦

Note.—Brackets connect conditions which *did not* differ significantly; a one-tailed Mann-Whitney U-test, 5% level of confidence is used throughout.

Thus, with regard to the relationship between *low-* and *intermediate-FU* on both rate of *IL* and on amount of *RI* generated, these results provide support for the reliability of the results already obtained in Experiment I.

Insofar as rate of learning may be assumed to vary directly and continuously with *FU*, the results of this experiment also provide some validation for the success of the attempt to make the *intermediate* level *IL* list in Experiment II of a higher order of *FU* than was that used in Experiment IB. The *intermediate* condition of Experiment IB resulted in 91.2% correct responses on the last (eighth) *IL* trial, while that of Experiment II resulted in 98.1% correct after the same amount of learning (compare Tables I and II). The difference between these two conditions was significant beyond the 5% level ($U = 50$, $n_1 = 12$, $n_2 = 14$), and in the required direction, thus supporting the notion that the *intermediate* condition of the second experiment was of a significantly higher level of *FU* than was that of the first experiment.

Discussion. The most important findings of the present investigation, as far as any general theory of learning and memory-interference is concerned, may be summarized as follows: (1) the observed *discontinuity* in magnitude of *RI,* when the complete sentence type of *IL* structure was compared with *any* lower order *IL* structure, no matter how similar they may be in their general "sentential" characteristics; (2) *the absence of any difference* between the intermediate (disrupted "sentence") and the lowest (random order) *IL* structures, no matter how different they may be in their "sentential" characteristics; and (3) the direct and *continuous* relationship between *FU* and *rate of IL,* over the entire range of *FU* sampled in these two experiments.

The theoretical problems which these findings pose for organizational theories, on the one hand, and associationistic theories, on the other, differ in some respects and in some respects overlap.

Discontinuity between Complete Sentences and Lower Order Structures

Within organizational theories, this *discontinuity* between the complete sentences and all lower order structures may be subsumed under the concept of *functional emergence—i.e.* the notion that after a certain level of organization has been achieved, the functional characteristics of the system as a whole become 'suddenly' (discontinuously) transformed. Within the present investigation, complete sentences may be thought of as different from all lower order structures in that 'a sentence is a group of words which expresses a *complete* thought.' The *completeness* of the thought expressed by the sentence may be regarded as the critical factor which places this level of *IL*-structuring *on a totally different continuum* from all lower *FU*-structures.

In the associationistic theories there does not appear to be any equivalent place for discontinuous functions *per se* and, indeed, the tendency has always been to try to reduce observed discontinuities in functional relationships to lower order continuous relationships. How such reduction can be achieved in the present case is not at all clear.

Absence of Difference between the Intermediate- and Low-FU Structures

The absence of any difference between the disrupted sentences (*intermediate-FU* condition in both experiments) and mere random orderings of items (the *low-FU* condition) in terms of amount of *RI* generated poses a serious theoretical problem for organizational theory. From this

point of view, one would have to concede that the *intermediate-FU* material in both experiments had a higher *potential* for organization (*i.e.* had greater *prägnanz*) than did the respective *low-FU* material. The differences in rate of *IL* for these two conditions clearly indicated that the *Ss* did, in fact, utilize this difference in organizational potential during the *IL*-phase of the task. Since greater organizational potential should imply greater boundary-strength for the memory-structures resulting from the learning of the *intermediate-FU* list, the items within this memory-structure should be less available to interfere with recall of *OL* during the *RL*-phase of the task.[5] The lack of difference between the *intermediate-* and the *low-FU* conditions in both experiments raises a serious question concerning the adequacy of this formulation.[6]

This same lack of difference, when taken in conjunction with the *rate of IL* data, raises a somewhat similar, and equally serious, problem for associationistic theories. Within this framework, the *strength of an associative bond* between successive items in a list should increase as a function of the *number of reinforced trials* which each link in the associative chain has previously received. It has generally been assumed that each correct response on a serial anticipation-task constitutes a 'reinforced trial' for the *S-R* pair for which that item is the response. Within the present context, in which exactly the same items appear in the *OL* and *IL* material, albeit in a different order, one would expect that a greater number of reinforced trials on the *IL* material should result in a net decrease in the strength of the *S-R* bonds formed during the learning of the *OL* material, *i.e.* a greater 'forgetting' of the *OL* with increased learning of the *IL*. Thus, according to this method of analysis, when *OL* is held constant, *RI* should vary directly as a function of number of reinforced trials. Both the absence of a difference between the *intermediate-* and the *low-FU* conditions, and the small magnitude of *RI* under the *high-FU* condition, seem incompatible with this set of assumptions.

While the author is fully cognizant of Melton and Irwin's findings that

---

[5] For a summary of this point of view, see Werner, The effect of boundary strength on interference and retention, this JOURNAL, 60, 1947, 598-599.

[6] Werner (*op cit.*) found a similar lack of difference in amount of proactive interference when previous learning of a *schematic* sentence structure was compared with previous learning of randomly selected words. Werner does not, however, report the actual *learning* data for the prior learning material, nor does he report any procedure in which *complete*, rather than merely *schematic* sentences were used for prior learning. It seems reasonable to assume, however, that the schematic sentences used in Werner's study correspond roughly to the *intermediate-FU* material of the present study, while his randomly selected list of words corresponds roughly to the *low-FU* material used here.

RI tends to decrease with *IL-overlearning*,[7] and while it might be argued that the *high-FU* condition resulted in a high degree of overlearning during *IL*, still (a) the magnitude of the difference between the *high-FU* condition and the two lower *FU* conditions in terms of amount of *RI* generated, far exceeds anything even hinted at by the data presented in the Melton and Irwin report, and (b) the line of reasoning employed by Melton and Irwin would then lead to the expectation that the *intermediate-FU* condition should, in some manner, differ from the *low-FU* condition in amount of *RI* generated, by virtue of its significantly greater number of reinforced *IL* responses.

Conclusions

Despite the difficulties encountered in attempting to explain the findings of these investigations within existing theoretical formulations of the mechanisms underlying learning and memory interference, the following empirical generalizations seem justified by the results.

When the internal structure of the *IL* material is varied along a dimension of *FU*, and when *FU* is manipulated by varying the degree to which the word order in the *IL* list approximates that of a complete and meaningful sentence, *then:*

(1) *Rate of IL* varies directly and continuously with *FU*; and (2) *amount of RI* varies inversely and discontinuously with *FU*, such that (a) the highest *FU*-list (word orders which form complete and meaningful sentences) result in significantly less *RI* than word orders which do not achieve this degree of organization, no matter how closely they may approach it, while (b) the lower *FU*-lists (disrupted sentences and random orders of words) did not differ significantly from each other, no matter how different they were in level of *FU* (as indexed by rate of *IL*).

While any speculation as to what kinds of theoretical assumptions or hypothetical mechanisms might be invoked to account for these results seems premature at this point, one empirical question seems worthy of immediate further investigation: What would the results have been if some mode of organization other than the "sentential" mode used here had been employed for varying the degree of *FU* of the *IL* material? What, for example, would be the effect upon *RI* if the *IL* material had been organized according to different types of rhythmic patterns rather than according to degree of sentential meaningfulness? The answer to

---

[7] A. W. Melton and J. M. Irwin, The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses, this Journal, 53, 1940, 173-203.

questions of this type might help in answering the more general question of whether the results obtained in this study are unique to the sentential mode of organization, *i.e.* whether there is something special about sentences, *per se,* or whether this represents some general characteristic of the dimension which has been referred to here as *FU*. This problem is currently under investigation.

Summary

Two experiments are reported in which we varied internal structure of the *IL*-material, in an experimental paradigm for retroactive inhibition, along a dimension referred to as *degree of functional unity.* It was hypothesized that increasing *FU* should make items within the *IL*-structure less available to interfere with *OL* during *RL,* hence that *RI* should vary inversely and continuously with degree of *FU* of the *IL* material.

The results, however, indicated that while (1) rate of *IL* varied directly and continuously with *FU,* (2) amount of *RI* varied inversely and discontinuously with *FU,* such that the highest *FU* condition resulted in the least *RI,* while the intermediate and lowest conditions resulted in more *RI* than the highest condition, and, at the same time did not differ from each other.

The difficulties posed by these findings for existing theories of learning and memory interference are discussed.

AMBIGUITY AND NEARNESS OF ANCHORS AS FACTORS IN ASSIMILATION

By GEORGE S. LARIMER, University of New Brunswick

According to the "assimilation-contrast model," there are three forms of assimilation that can occur when anchors do not differ greatly from the individual's preëxisting scale of judgment.[1]

(1) The first form of assimilation is characterized by a shift in the judgmental scale toward the value of the anchor. Typically, this form is found in psychophysical studies.[2] (2) The second form is described as follows: When the communication (anchor) is susceptible to alternative interpretations, *S*s with positions slightly removed from that of the communication may judge it more like their own position than it actually is.[3] This form of assimilation has been observed chiefly in studies of the effect of propaganda on attitudes, *e.g.* on perception of the movie, *Mr. Bigot*;[4] on the judgments of pro, neutral, and anti-fraternity messages; on judgments of pro, anti, and ambiguous statements about science;[5] and on neutral *S*s' judgments of the propaganda for prohibition in a dry state.[6] The effect has also been reported in judgments of brightness.[7] (3) The third form of assimilation, an enlargement of the scale, has been reported in studies dealing with changes in latitudes of acceptance of a social issue. Sherif and Hovland, in a study on attitudinal changes conducted during the 1956 Presidential campaign, reported that, for *S*s who held moderate positions, changes in latitudes of acceptance were "largely confined to the

* Received for prior publication May 18, 1965. This is a portion of a dissertation submitted to the Psychology Department, University of Utah, in partial fulfillment of a Doctor of Philosophy degree. The research was done under the co-sponsorship of Drs. Muzafer Sherif and B. Jack White.

[1] Muzafer Sherif and C. I. Hovland, *Social Judgment*, 1961, 177-207.

[2] Spaulding Rogers, The anchoring of absolute judgments, *Arch. Psychol.*, 37, 1941 (No. 261), 1-42; Muzafer Sherif, Daniel Taub, and C. I. Hovland, Assimilation and contrast effects of anchoring stimuli on judgment. *J. exp. Psychol.*, 55, 1958, 150-155.

[3] Sherif and Hovland, *op. cit.*, 149 ff.

[4] Eunice Cooper and Marie Jahoda, The evasion of propaganda: How prejudiced people respond to anti-prejudice propaganda, *J. Psychol.*, 23, 1947, 15-26.

[5] Melvin Manis, The interpretation of opinion statements as a function of recipient attitude, *J. abnorm. soc. Psychol.*, 60, 1960, 340-344; The interpretation of opinion statements as a function of recipient attitude, *ibid.*, 63, 1961, 76-81; P. G. Zimbardo, Verbal ambiguity and judgmental distortion, *Psychol. Reports*, 6, 1960, 57-58.

[6] Bernard Berelson, P. F. Lazarsfeld, and W. N. McPhee, Political perception, in Eleanor Maccoby, T. M. Newcomb, and E. L. Hartley (eds.), *Readings in Social Psychology*, 1958, 72-84; C. I. Hovland, O. J. Harvey, and Muzafer Sherif, Assimilation and contrast effects in reactions to communication and attitude change. *J. abnorm. soc. Psychol.*, 55, 1957, 244-252.

[7] S. I. Perloe, Assimilation as a consequence of categorization, *J. Pers.*, 29, 1961, 148-166.

addition . . . of a single position."[8] Thus, the results indicated an enlargement of the judgmental scale through the addition of new categories. Similar results were reported by Hovland, Harvey and Sherif in a study of the reactions in a dry state to persuasive communications about prohibition.[9]

The studies mentioned above on judgments of social issues seem to indicate that assimilation is dependent upon the ambiguity of a position advocated, as well as the magnitude of its discrepancy from a person's stand on the issue. Because most of our knowledge concerning the assimilative process has come from studies of the judgments of complex, affect-laden verbal material (*e.g.* attitudes), quantification of the effect has been difficult, and in some instances, impossible. The difficulty in quantification has led several investigators to generalize the results of studies of anchoring to the more complex results of studies of social judgments.[10] When, however, the results of anchoring studies[11] are compared with those involving judgments of social issues,[12] a conceptual gap appears, because the findings indicate different *kinds* of assimilation. The present study is an attempt to narrow this gap by demonstrating that judgments of physical distance can produce the same kinds of assimilation found in studies on judgments of social issues.

To do so, comparisons were made between the forms of assimilation produced by anchors which varied in ambiguity and in nearness to a judgmental scale for distance.

From the assimilation-contrast model and results of the studies summarized, two hypotheses emerge: (1) When the anchor is close to the judgmental scale, is judged, and is relatively unambiguous, the anchor will be defined in a new category, adjacent to the original judgmental scale (*i.e.* the original scale will be enlarged). (2) When an anchor close to the judgmental scale is judged, the judgments are relatively ambiguous; judgments of the anchor shift in the direction of the original judgmental scale.

Method. The design tested hypothesized effects on judgments using a graded stimulus-series. Anchors varied both in distance from the terminal stimuli of the series and in degree of ambiguity. The experiment consisted of three parts: a criterion, a standardization, and an anchoring session.

Subjects. Sixty-five men, undergraduate volunteers from the University of Oklahoma, served in all parts of the experiment.

---

[8] Sherif and Hovland, *op. cit.,* 166-167.
[9] Hovland, Harvey, and Sherif, *op. cit.,* 244 f.
[10] Sherif and Hovland, *op. cit.,* 177-207; Harry Helson, *Adaptation Level Theory,* 1964, 648-657.
[11] Rogers, *op. cit.,* 1-42; Sherif, Taub, and Hovland, *op. cit.,* 150-155.
[12] Berelson, Lazarsfeld, and McPhee, *op. cit.,* 72-84.

Stimuli. The stimuli were distances between two pin-points of light in a darkroom. The original stimulus-series consisted of a psychologically equal, horizontal distances of 7⅜, 8-19/32, 10-1/16, 11¾, and 13-9/16 in., hereafter referred to as 1 thorugh 5, respectively. A distance of 14-11/32 in. was used for the 'near' anchor; one of 17-1/16 in. for the 'medium' anchor.

Apparatus. Eight dial light assemblies were mounted horizontally on a panel 16 ft. 8 in. in front of *S*. The caps of the dial light assemblies were opaque, except for 0.025 in. holes drilled in their centers.

The apparatus automatically performed the following functions: (1) presented distances in a pre-programmed random sequence; (2) set the time-interval during which the lights remained on; (3) established a timed rest-interval after 25 distances had been presented; and (4) provided for simultaneous or successive presentations of the anchor-lights.

Procedure in the criterion sessions. Each *S* was seated before the apparatus in a darkroom and received taped instructions explaining the general purpose of the experiment and instructing him to "judge the distance, in whole inches, between two lights which will come on either simultaneously or successively"; that a rest-interval would occur after each series of 25 judgments; and that 25 practice trials would be given him. Then the automatic operation of the apparatus was explained.

Following the 25 practice trials, 100 judgments (20 judgments of each distance) of the original series were obtained, where distances were presented in a random order except for the restrictions that each distance was presented an equal number of times and that no distance was presented twice in succession. For each *S*, a median judgment of each distance of the original series was calculated, and those *Ss* unable to discriminate between stimuli were eliminated from the remaining two sessions of the experiment. The criterion for lack of discrimination was *S*'s inability to place the five stimuli into five categories whose medians were no less than one inch apart. (The criterion insured that any change in judgments in the anchoring sessions was not caused by lack of discrimination.)

Procedure. The five groups of 13 *Ss*, each of which participated in these sessions of the experiment, may be described, respectively, as (I) the ambiguous-near anchor (*A-NA*); (II) the unambiguous-near anchor (*UA-NA*); (III) the ambiguous-medium anchor (*A-MA*); (IV) the unambiguous-medium anchor (*UA-MA*); and (V) the control (*C*) groups. The groups were exposed to the standardization and anchoring-sessions approximately one week after completing the criterion-session. Procedure in the standardization-session was the same as in the criterion-session. In the anchoring session, the distances of the original series were presented in the same random order as presented in the criterion-session, but the anchor-distance was presented before each distance of the original series. Thus, a variation of the method of constant stimuli was used, the variation being that the anchor was judged each time that it appeared. The ambiguity of the anchor was varied by simultaneous or successive presentation of the anchor-lights. In the relatively unambiguous condition, the two lights came on simultaneously; in the ambiguous condition, the two lights came on successively with 4 sec. between exposures.

At the beginning of the standardization-session, each *S* was again seated in the darkroom before the apparatus and given pre-recorded instructions which stated

that the session was in two parts and that the procedure was the same as before. Each S then made 100 judgments of the original series. The anchoring session of 200 judgments (100 of the original series and 100 of the anchor) followed the standardization-session after a 15-sec. rest-interval with no prior knowledge or additional instructions.

The four experimental groups were distinguished as follows: *S*s in the *A-NA* and *UA-NA* groups (Groups I and II) made judgments of the original series paired with the near anchor, presented in the successive and simultaneous modes

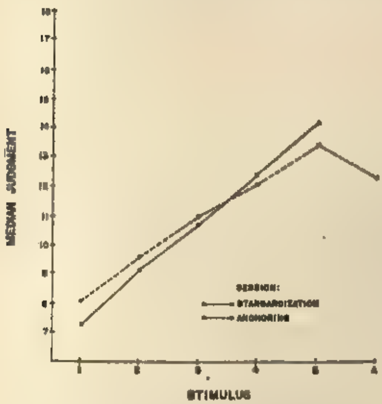

FIG. 1. MEDIAN JUDGMENTS OF GROUP I
(The anchor in this and the following Figures is denoted by "A" on the abscissa.)

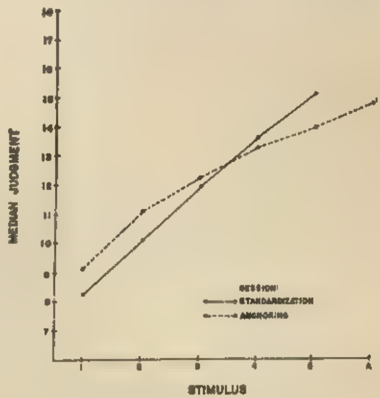

FIG. 2. MEDIAN JUDGMENTS OF GROUP II

respectively. The *S*s in the *A-MA* and *UA-MA* groups (Groups III and IV) made judgments of the original series paired with the medium anchor, presented in the successive and simultaneous modes, respectively. In contrast to these *S*s, those in the C group (Group V) made 100 judgments of the original series wtihout the presence of an anchor, to assess the cumulative effects of practice judging the original series.

Results. An over-all picture of the results is presented in Figs. 1, 2, 3, 4 and 5. Comparisons of the judgments from the standardized session with those in the sessions with anchors reveal that all groups, including Group C, used fewer categories in the anchored than in the standardized sessions. This constriction of the judgmental scale in all groups is evidenced by the decrease in verticality of the lines representing median judgments in the anchoring-session.

Because there is heterogeneity of variance and non-normality of data

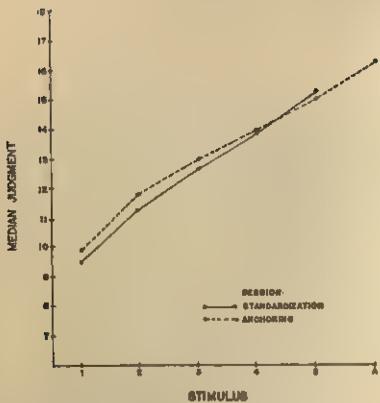

Fig. 3. Median Judgments of Group III

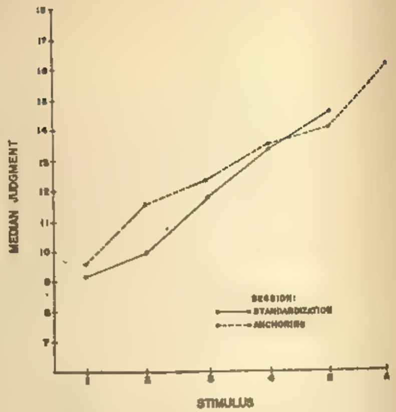

Fig. 4. Median Judgments of Group IV

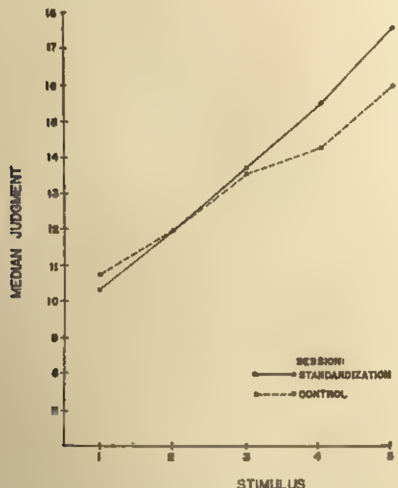

Fig. 5. Median Judgments of Group V
(The control group with which no anchor was experienced.)

from studies of anchoring effects, Salzinger recommends the use of non-parametric statistics for analysis.[13] Accordingly, all statistical comparisons in the present study were made with non-parametric techniques. Five change-scores were obtained for each of the five groups by subtracting their grand median judgment of each stimulus in the standardized sessions from their grand median judgment of each stimulus in the original series during the anchoring-session. The purpose of obtaining the change-scores was to assess the effects of the anchors on judgments of the original series. Results of a comparison of change-scores between groups, using the Kruskal-Wallace one-way analysis of variance indicated that the five groups did not differ significantly in the amount of change exhibited ($H = 4.34$; $df = 4$; $p > 0.30$).[14] Since the groups did not differ, it can

TABLE I

COMPARISON OF SCORES (DIFFERENCES) FOR THE EXPERIMENTAL GROUPS

(Comparisons were made with the Wilcoxon matched-pairs signed-ranks test. Since there were no cases where difference-scores were zero, $N = 13$ in each group.)

Group	Direction	T	p†
I(*A-NA*)	A<5*	17	<.05
II(*UA-NA*)	A>5	0	<.01
III(*A-MA*)	A=5	19	>.05
IV(*UA-MA*)	A>5	0	<.01

* 'A' = anchor; '5' = Stimulus 5.
† Two-tailed probability values.

be concluded that changes which did occur for all groups were a function of practice judging the original series and not effects of the anchor.

Analysis of judgments of the anchor. A difference-score was obtained for each *S* in the four experimental groups by subtracting his median judgment of the anchor from his median judgment of Stimulus 5 in the anchoring-session. A positive difference-score indicated assimilation of the anchor; a negative difference-score indicated that the anchor was defined as falling outside the judgmental scale originally used for the series.

A comparison of each group's median judgment of Stimulus 5 with its median judgment of the anchor was made with the Wilcoxon Matched-Pairs Signed-Ranks test.[15] The results of the comparisons are presented in Table I.

---

[13] Kurt Salzinger, Techniques for computing shift in a scale of absolute judgment, *Psychol. Bull.*, 53, 1956, 394-401.
[14] Sidney Siegel, *Nonparametric Statistics*, 1956, 184-193.
[15] Siegel, *op. cit.*, 75-83.

As Table I shows, the groups which experienced the ambiguous anchors significantly displaced their judgments of the distance of the anchor toward the terminal category of their scale. Group I (the *A-NA* group) judged the anchor as being significantly shorter than Stimulus 5; in fact, it judged the anchor equal to Stimulus 4 ($T = 38$; $N = 13$; $p > 0.05$). Group III (the *A-MA* group) judged their anchor as equal to Stimulus 5. Since both of the ambiguous anchor-groups defined their anchors as being within their judgmental scales, the second hypothesis that judgments of ambiguous anchors will shift in the direction of the original scale was supported. In addition, since the unambiguous anchor-groups defined their anchors as being outside their judgmental scales, indicating the addition of new categories to the scale, the first hypothesis that unambiguous anchors will be defined in new categories adjacent to the scale was also supported.

TABLE II

COMPARISONS OF SCORES OF ALL EXPERIMENTAL GROUPS
(Comparisons were made with the Mann-Whitney U test.)

Group	U	p*	Group	U	p*
I(*A-NA*)>II(*UA-NA*)	23	<.002	II(*UA-NA*)>III(*UA-MA*)	3	<.002
III(*A-MA*)>IV(*UA-MA*)	37	<.02	I(*A-NA*)>IV(*UA-MA*)	1	<.002
I(*A-NA*)>III(*A-MA*)	30	<.02	IV(*UA-NA*)=III(*A-MA*)	78	>.05

* Two-tailed probability values.

To test these hypotheses further, a comparison of difference-scores between judgments of the anchor and Stimulus 5 in the anchoring-session was made with the Kruskal-Wallace non-parametric analysis of variance. This test revealed that the average difference-scores for the four experimental groups were not equal ($H = 41.81$; $df = 3$; $p < 0.001$). The results of Mann-Whitney U tests comparing the difference-scores for each group with those of all other groups is shown in Table II.[16]

Perhaps the most significant finding in these comparisons was that Group II and III (the *UA-NA* and *A-MA* groups) did not differ significantly in difference-scores even though objectively their respective anchors were grossly different. This result further supports the second hypothesis that the greater the ambiguity of the anchor the more will judgments of the anchor be displaced in the direction of the scale.

Discussion. The control group changed their judgments of the original series as much as the experimental groups, indicating that change was produced simply by practice. Previous studies dealing with anchoring effects seldom included control groups. Results of the present study clearly

[16] Siegel, *op. cit.*, 116-127.

indicate, however, that studies of assimilation-effects (effects which tend to be subtle, by comparison with contrast effects), should employ a control group to differentiate effects of practice in judging stimuli from effects of the anchor.

Both near and medium ambiguous anchors were displaced toward the scale, whereas unambiguous anchors were defined in new categories. Distance of the anchor in the present study was, however, based only on two points. Still open to question is just how far the ambiguous anchor could be removed before it ceases to be displaced toward the scale and begins to be displaced away from it. The results of the 1956 Presidential election, reported by Sherif and Hovland, indicate that at some point, when the ambiguous anchor is removed too far from latitude of acceptance, displacement of the anchor will be away from the scale (contrast effect).[17] These studies, in contrast to anchoring studies, used highly ego-involved *S*s and ego-involving anchors. It may be, as Sherif and Hovland state, that latitude of acceptance is quite broad with non-ego-involved scales and non-ego-involved anchors, so long as the anchor is ambiguous.[18] Hence, a non-ego-involving anchor which is ambiguous might never be defined as more discrepant from an *S*'s judgmental scale than a more unambiguous anchor.

The results of the present study suggest that judgmental scales for physical stimuli have broad assimilative ranges for ambiguous anchors and a wide area where unambiguous anchors are defined in new categories. This generalization, however, assumes that the judgmental scale is formed with unlimited categories available to *S* (in other words, with *S* deriving his own scale rather than having it imposed by *E*) and that judgments of the anchor are allowed. If the above assumptions are not met, the assimilative range may be quite different. When the above assumptions are met, the present study demonstrates that some of the assimilative effects found in judgments of affect-laden verbal material can be duplicated in judgments of simple, physical stimuli.

Results supported the second hypothesis that ambiguous anchors would be displaced in the direction of the original scale. Such results lend credence to the adage, "Some propagandists . . . believe that ambiguity often promotes effectiveness, since each subject is then free to define the matter in terms satisfactory to himself."[19] If, however, by 'effectiveness' the propagandists mean that an ambiguous anchor produces more attitude-

---

[17] Sherif and Hovland, *op. cit.,* 177-207.
[18] *Idem,* 177 f.
[19] Berelson, Lazarsfeld, and McPhee, *op. cit.,* 73.

change than an unambiguous anchor, then the results of the present study do not support such a conclusion. In the present study, none of the anchors produced a significantly greater change in judgment of the original series than could be attributed to practice. Thus, the addition of new categories to the scale for the unambiguous anchor was the only change exhibited in the judgmental scales. On the basis of evidence from the present study, it appears that the sort of assimilation described by Berelson *et al.* would not bring about much, if any, attitude-change. This conclusion is supported by Cooper and Jahoda, who found that distorting a message to fit one's preconceived prejudices was one form of propaganda evasion, *i.e.* one technique the prejudiced person uses to prevent change in his prejudiced attitudes.[20]

From the above results, it may be concluded that ambiguous anchors are effective only when the propagandist wishes to leave a favorable impression with his audience. If he wishes to *change* attitudes, perhaps a better approach is to attempt to add new categories to the audience's latitudes of acceptance, by using a relatively structured communication which is not far removed from the audience's current latitudes of acceptance.

Summary

Five groups of *S*s judged a series of distances between lights under standardized and variable conditions, *i.e.* with the introduction of an anchor. Under the standard condition the *S*s were allowed to establish their own scales. Under the variable condition, Group I judged the original series in the presence of an ambiguously near anchor (*A-NA*); Group II, in the presence of an unambiguously near anchor (*UA-NA*); Group III, in the presence of an ambiguously medium anchor (*A-MA*); and Group IV, in the presence of an unambiguously medium anchor (*UA-MA*). In all cases the anchor was also judged. Group V, the control group (*C*), judged the series without the introduction of an anchor. Significant displacements occurred in the judgments of Groups I and II (the *MA-NA* and *A-MA* groups). Groups II and IV (the *UA-NA* and *UA-MA* groups) enlarged their judgmental scales by adding new categories adjacent to their original scale to include the anchors. Practice was found to account for changes in judgments of the original series.

---

[20] Cooper and Jahoda, *op. cit.*, 15-26.

EXPERIMENTAL COMPARISON OF CLASSICAL AND INSTRUMENTAL APPETITIVE CONDITIONING

By RONALD G. WEISMAN, Queens University

Perhaps the ideal method for assessing similarities and differences in the effects of instrumental and classical conditioning would be to train the same responses to the same stimuli using a yoked subject technique. Such experiments have been performed with aversive stimuli,[1] but the reinforcement is probably not the same in classical and instrumental aversive conditioning. In the classical paradigm, the onset, and in the instrumental paradigm, the offset, of the aversive stimulus is the reinforcement.[2]

Since it is generally held that the onset of appetitive stimuli serve as reinforcement in both conditioning paradigms, water-reinforcement was used in the present research. This choice was dictated also by evidence that the licking response can be modified by both classical[3] and instrumental[4] procedures. Specifically, the present study reports the results of comparisons of a modification of the classical technique described by Debold, Miller, and Jensen[5] and a discrete-trials instrumental procedure having similar temporal parameters.

EXPERIMENT I

Experiment I was designed to establish that the noncontingent procedure produces classical conditioning rather than pseudoconditioning or nondiscriminated response-modification. Toward that end, the effects of three procedures were compared: presentation of the *US* alone, pseudo-

---

* Received for publication August 6, 1964. This study is adapted from a Doctoral dissertation submitted to Michigan State University.

[1] Harold Schlosberg, Conditioned responses in the white rat: II. Conditioned responses based upon shock to the foreleg, *J. genet. Psychol.*, 49, 1936, 107-138; W. J. Brogden, E. A. Lipman, and Elmer Culler, The role incentive in conditioning and extinction, this JOURNAL, 51, 1938, 109-117; J. W. Moore and I. Gormezano, Yoked comparisons of instrumental and classical eyelid conditioning, *J. exp. Psychol.*, 62, 1961, 552-559; Robert Fromer, Conditioned vasomotor responses in the rabbit, *J. comp. physiol. Psychol.*, 56, 1963, 1050-1055.

[2] G. A. Kimble, *Hilgard and Marquis' Conditioning and Learning*, 1961, 79-80.

[3] R. C. Debold, N. E. Miller, and Donald Jensen, Effect of strength of drive determined by a new technique for appetitive classical conditioning of rats, *J. comp. physiol. Psychol.*, 59, 1965, 102-108.

[4] L. B. Wyckoff, Joseph Sidowski, and D. J. Chambliss, An experimental study of the relationship between secondary reinforcing and cue effects of a stimulus, *J. comp. physiol. Psychol.*, 51, 1958, 103-109.

[5] Debold, Miller, and Jensen, *op. cit.*, 103-105.

conditioning (unpaired presentation of *CS* and *US*), and classical conditioning.

Subjects. The *Ss* were eight experimentally naïve male gray hooded rats of the Long-Evans strain. The *Ss* had access to water for 30 min. a day for three weeks prior to the experiment. During the experiment, they had access to water for only 5 min. after each daily session. A given *S* was run at approximately the same time each day.

Apparatus. The apparatus consisted of two matched systems. Each experimental box was 8 in. long, 3½ in. wide, and 4 in. high, and was constructed of Plexiglas with a ⅛-in. stainless-steel grid in the floor. At one end, a 1½-in. enclosed drinking well extended out of the box. Water was presented in small quantities by a solenoid-operated valve from a #11 hypodermic needle (ground flat and smooth) which projected through a small opening at the bottom of the well. A small copper ring encircled this opening to prevent gnawing and to measure licking.

The *CS* or S^D was a 10-w. (50-55 ft.c.) lamp mounted on the outside of the box next to the well. Licking was measured by contact-relays and recorded on an event-recorder and on counters. Each of the two boxes was enclosed in a converted refrigerator-shell containing an exhaust fan (60 cu. ft. per min.) and a white-noise speaker (50 db.). The general level of illumination in the box was 5-10 ft.c. (depending on *S*'s position). Electronic timers and a punch-tape timer provided all temporal intervals.

Procedure. All *Ss* had four sessions during which 40 *US* (water) presentations with a 90-sec. variable intertrial interval were given. The first session served as pretraining to the apparatus and water-delivery system. During the remaining three sessions, responses were recorded during the 3-sec. period just preceding water-presentation. After the three '*US*-only' sessions, an extinction-session was given.

The *Ss* were then divided randomly into two groups of four *Ss* each. Group *A* received eight days of classical conditioning (40 trials per day). The interval between light-onset and water-presentation (*CS-US* interval) was 3 sec., and the light remained on for an additional 2 sec. (an overlapping *CS*). The intertrial interval was variable with a mean of 90 sec. and a range of 60-120 sec. Group *B* received eight days of a pseudoconditioning procedure. That is, water and light were presented 40 times per session in a random sequence with the restriction that neither followed itself more than twice and that water and light were separated by at least 30 sec. For each stimulus, independently, there was a 90-sec. variable intertrial interval. The groups received an extinction session following pseudoconditioning or classical conditioning. Both groups then were classically conditioned and, afterward, extinguished.

Results. Two response-measures were considered: one (*CR*) was the percentage of trials on which licking occurred during the 3-sec. period in which the light was presented but on which there were no licks in the 3-sec. period immediately preceding the light; the second (pre-*CR*) was the percentage of trials on which licking occurred both in the 3 sec. before the *CS* and during the *CS-US* interval.

Fig. 1 presents the results for two representative *Ss* in Group *A*. The

curves show relatively smooth acquisition and reconditioning with minimal pre-*CR* activity. The results for two sample *S*s in Group *B* are presented in Fig. 2. As in Group *A* the level of pre-*CR* responding was low, as was the level of response in the training with *US* alone, but the

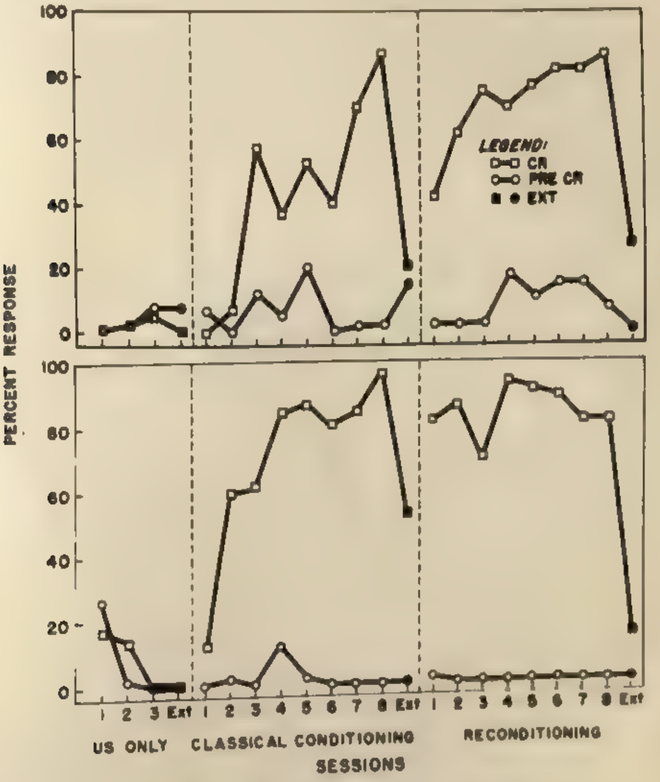

FIG. 1. THE PERCENTAGE *CR* AND PRE-*CR* RESPONSE FOR TWO SAMPLE *S*S IN GROUP *A* (EXPERIMENT I)

pseudoconditioning procedure did not increase the number of *CR*s over the *US*-only level. When the classical conditioning procedure was introduced, the *CR*-level increased markedly in three of the four *S*s. The fourth *S* (*B*-4) which had shown the characteristically low rate of response in pseudoconditioning, showed in classical conditioning a high pre-*CR* level. A decrease in the length of the intertrial interval failed to improve the situation, but, when the intertrial interval was lengthened to 120 sec.,

the level of pre-*CR* decreased rapidly to the same level as the other *Ss* in Group *B*.

Prior pseudoconditioning sessions may account for the relatively slow

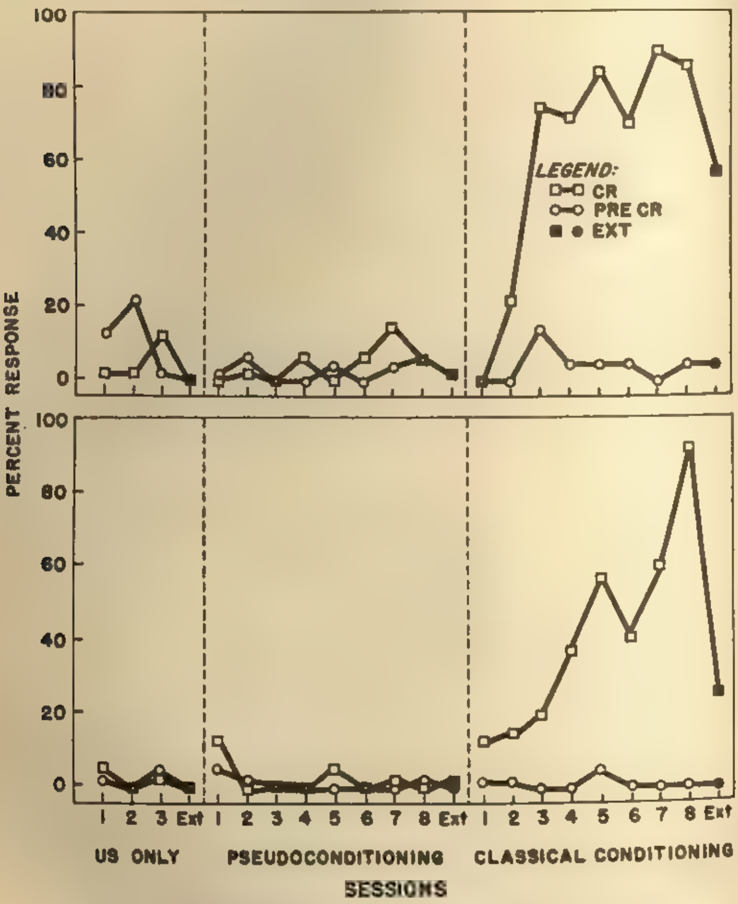

FIG. 2. THE PERCENTAGE *CR* AND PRE-*CR* RESPONSE FOR TWO SAMPLE *Ss* IN GROUP *B* (EXPERIMENT *II*)

rate of conditioning and high frequency of pre-*CR* in *B*-4. Since *CS* and *US* are never paired in pseudoconditioning, it can also be interpreted as inhibitory training. Thus, the *Ss* in Group *B* may have learned *not* to respond to the *CS* during pseudoconditioning.

Another explanation of the high pre-*CR* rate in *B*-4 is that responding was adventitiously reinforced. Operationally, classical conditioning and

adventitiously reinforced operant behaviour (superstition) are quite similar,[6] but there are some important differences in the effects of these procedures. Operant discriminations usually involve the extinction of unreinforced responding in the S$^{\Delta}$, while the level of intertrial responding in classical conditioning usually differs little from its original level. Of the eight *Ss* conditioned in Experiment I, and the many *Ss* used in pilot work and other experiments, only *B*-4 had any appreciable rate of pre-*CR* to be extinguished. Furthermore, when 'superstitions' develop without explicit instrumental training, as in Skinner's study,[7] the response-topography tends to be unstable, often changing over sessions or in reconditioning,[8] in marked contrast to classical conditioning where stability

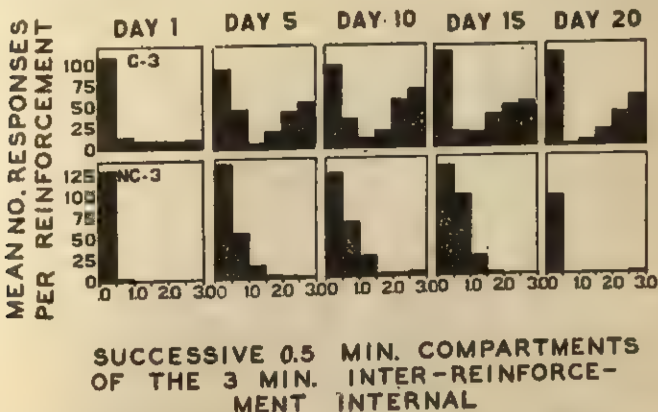

FIG. 3. THE TEMPORAL DISTRIBUTION OF RESPONSES DURING THE 3-MIN. INTER-REINFORCEMENT INTERVAL FOR TWO SAMPLE *Ss* (EXPERIMENT II)

of response is the rule. The topography of the lick-response in Experiment I was quite stable both from *S* to *S* and over sessions. Thus, the results suggested that classical conditioning rather than adventitiously reinforced operant conditioning was involved.

EXPERIMENT II

The conclusion that the noncontingent procedure of Experiment I is a form of classical conditioning rather than superstition was based in part on the species-specific, as opposed to the subject-specific, nature of the

---

[6] Kimble, *op. cit.*, 197.
[7] B. F. Skinner, Superstition in the pigeon, *J. exp. Psychol.*, 38, 1948, 168-172.
[8] Murray Sidman, *Tactics of Scientific Research*, 1960, 349-355.

topography of the conditioned response. Not all investigators developing new techniques of classical conditioning deal with responses as easily defined as licking or salivation.[9] In these cases, the problem of operationally differentiating operant and classical conditioning procedures has been more acute and of considerable theoretical interest.[10]

Experiment II was designed to differentiate instrumental conditioning procedures including adventitious reinforcement from the classical paradigm with respect to the effects of interstimulus-interval. Two procedures were used: temporal conditioning, and delayed conditioning with an extroceptive stimulus.

Operant conditioning in the rat with interreinforcement intervals of 3 min. or more (*FI*-3 min.) is common, while Pavlovian temporal conditioning (no *CS*) in the rat using intervals of this length is rare indeed. Similarly, with operant behavior, including superstition, a change from *FI*-15 sec. to *FI*-60 sec. hardly disrupts performance at all,[11] while in the Pavlovian paradigm a precipitious jump from a 15-sec. *CS-US* interval to a 60-sec. *CS-US* interval is usually quite disruptive.[12] Thus, it was expected that *S*s under the classical procedure would be severely affected by long temporal intervals, while *S*s under the operant and superstitious-operant procedures would show little if any decrement in performance.

Most recent studies of adventitious reinforcement have taken advantage of past experimental histories of instrumental conditioning. In Experiment II, the effects of an adventitious procedure were assessed using two *S*s with extensive experimental histories of instrumental lick-conditioning. These *S*s were shifted to the noncontingent procedure used for the Pavlovian *S*s during Experiment II. Like the 'regular' instrumental *S*s, but unlike the *S*s without instrumental histories (Pavlovian *S*s), these *S*s were expected to show no decremental effect due to interstimulus-interval.

Subjects. Eight male gray rats of the Long-Evans strain were used. Four of the *S*s (*C*-1, *C*-2, *NC*-1, and *NC*-2) were experimentally naïve. Two *S*s (*C*-3 and *NC*-3) were from Group *A* of Experiment I. Two *S*s (*C/NC*-1 and *C/NC*-2) had about 40 sessions of instrumental training previous to Experiment II. The conditions of deprivation were the same as in Experiment I.

Apparatus. The experimental boxes, equipment for programming, and the recorder of Experiment I were used, along with a print-out counter that recorded the temporal distribution of responses.

Procedure. On the first two days of training *US*-only pretraining was given to all *S*s. They then were matched on the basis of consummatory licking. One *S* from each pair was randomly assigned to each condition: instrumental or contingent (*C*), and classical or noncontingent (*NC*) reinforcement.

During the first phase of Experiment II proper, contingent *S*s had 20 daily ses-

---

[9] Nicholas Longo, Suzanne Klempay, and M. E. Bitterman, Classical appetitive conditioning in the pigeon. *Psychon. Sci.*, 1, 1964, 19-20; S. S. Pliskoff, T. D. Hawkins, and J. E. Wright, Some observations on the discriminative stimulus hypothesis and rewarding electrical stimulation of the brain, *Psychol. Rec.*, 14, 1964, 179-184.

[10] F. S. Keller and W. N. Schoenfeld, *Principles of Psychology,* 1950, 63-64; Kimble, *op. cit.*, 98-101.

[11] Skinner, *op. cit.*, 171.

[12] I. P. Pavlov, *Conditioned Reflexes*, 1927, 89.

sions of 25-30 reinforcements each on *FI*-3 min. Noncontingent *Ss* received reinforcements at the same time, but independently of response (yoked-*S* technique).

In the second phase of Experiment II, the contingent *Ss* were shifted to a discrete-trials procedure. Each trial began with the onset of the signal light, which remained on for 6 sec. independently of response. If *S* licked the dry water-source between the beginning of the third second and the end of the sixth second of a trial, a reinforcement (drop of water) was delivered. Responses during the first 3 sec. of a trial or during the intertrialinterval were not reinforced. The intertrial interval varied from 150-210 sec. with a mean of 180 sec. Responding during the delay-period of each

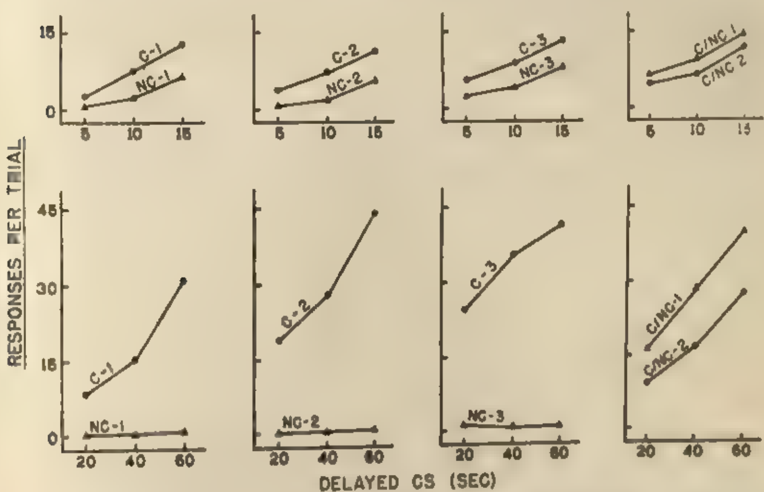

FIG. 4. MEAN AMOUNT *CR* PER THIRD OF THE INTER-STIMULUS INTERVAL (EXPERIMENT II)

C/NS-1 and *C/NC*-2 had extensive histories of instrumental licking prior to the Experiment.

trial was quite predictable, since the procedure was in all respects identical to the *FI*-component of a multiple schedule.[12] The noncontingent *Ss* were again yoked with the contingent *Ss* (*i.e.* the contingent *S* produced reinforcement for both *Ss*). In effect, the noncontingent *Ss* had training with a 3-sec. *CS-US* interval, a 6-sec. overlapping *CS*, and variable intertrial interval with a mean of 180 sec. The purpose of this rather involved procedure, then, was to equate the instrumentally and classically conditioned *Ss* with respect to a number of temporal variables (*e.g.* effective *CS-US* interval, *CS*-duration, *CS*-offset, and intertrial interval).

In the second phase of Experiment II, two *Ss* (*C/NC*-1 and *C/NC*-2) were shifted back to the noncontingent procedure after extensive training under the contingent procedure. Any irreversible effects or superstition due to instrumental training would be reflected in the results for these *Ss*.

[12] C. B. Ferster and B. F. Skinner, *Schedules of Reinforcement*, 1957, 527-529.

The interval between light-onset and reinforcement (*FI*-component for contingent *Ss*, or *CS-US* interval for noncontingent *Ss*) was investigated. All *Ss* had six sessions with the 3-sec. interval to stabilize the *CR*, followed by two sessions each with 5- and 10-sec. intervals, and three sessions each with the 15- and 60-sec. intervals. All *Ss* had 40 trials per daily session.

Results. Fig. 4 shows the mean number of responses in successive half-minutes of the 3-min. inter-reinforcement interval over the 1st, 5th, 10th, 15th, and 20th sessions of the first phase of Experiment II. Because of the similarity in results, only the data for *C*-3 and *NC*-3 are shown. All the contingent *Ss* (*C*-1, *C*-2, and *C*-3) developed typical *FI*- responding with the consummatory response predominating only in the earliest compartments. By contrast, the noncontingent *Ss* (*NC*-1, *NC*-2, *NC*-3) gave no evidence of temporal conditioning over the 20 sessions; except for consummatory responses, inter-reinforcement responding was negligible.

TABLE I

PERCENTAGE OF CONDITIONED RESPONSES IN THE FINAL SESSION AT EACH *CS-US* INTERVAL IN EXPERIMENT II

S	3 sec.	5 sec.	10 sec.	15 sec.	60 sec.
C-1	97.5	100.0	100.0	97.5	100.0
NC-1	97.5	90.0	92.5	90.0	30.0
C-2	95.0	95.0	100.0	97.5	97.5
NC-2	95.0	95.0	90.0	97.5	45.0
C-3	100.0	100.0	100.0	100.0	97.5
NC-3	95.0	97.5	95.0	95.0	55.0
C/NC-1	95.0	97.5	90.0	90.0	95.0
C/NC-2	95.0	100.0	95.0	92.5	95.0

Table I gives the percentage of *CRs* for each *S* during the final session at each *CS-US* interval in the second phase of Experiment II. All *Ss* showed excellent maintenance of the conditioned response at intervals of 3, 5, 10, and 15 sec., but only the contingent and previously contingent *Ss* continued at this level with the 60-sec. interval.

An analysis of amount of *CR* (number of licks per trial) in successive thirds of the 15- and 60-sec. *CS-US* intervals is presented in Fig. 5. At the 15-sec. interval, all *Ss* showed inhibition of delay or *FI*- scalloping. At the 60-sec. interval, this pattern was maintained at an even higher level in the *C* and *C/NC Ss*, but responding was negligible, in the *NC Ss*.

The effect of the instrumental procedure on the *C* and *C/NC Ss* was to facilitate responding at all intervals, but the effect was most pronounced at the 60-sec. interval. The *NC Ss* did not, however, maintain the response at the 60-sec. interval. It is clear that the *C* and *C/NC* procedures repre-

sent two kinds of instrumental training: operant and superstitious operant, while the *NC* procedure represents classical or Pavlovian training. These results suggest that it is possible to distinguish between instrumental and classical conditioning procedures without recourse to supposed differences in the *UR*.

Summary

Two experiments with the conditioned lick-response were reported. Experiment I demonstrated that pairing of a *CS* and water could result in classical conditioning as opposed to pseudoconditioning. *S*s trained by the classical procedure in Experiment II made fewer responses to the *CS*, especially with a long CS-US interval (60 sec.), than *S*s trained by the instrumental procedure. Reinforcement contingent upon the first response made 3 min. after the last reinforcement (*FI*-3 min.) resulted in typical *FI*-scalloping, while reinforcement every 3 min. independently of response did not produce temporal conditioning. These results suggest that the two conditioning paradigms differ in terms of sensitivity to the length of the interstimulus-interval.

RECOVERY OF FOVEAL ACUITY FOLLOWING EXPOSURE TO VARIOUS INTENSITIES AND DURATIONS OF LIGHT

By Jo Ann S. Kinney and Mary M. Connors, Groton, Connecticut

In any situation in which the eye is used at dim levels of illumination, the effect of a brief exposure of bright light on the dark-adapted fovea has many practical consequences. The usual tasks performed in such a situation involve foveal acuity, as the individual turns his attention and his fixation toward the target that must be resolved. Despite its importance, few data are concerned with this problem of the recovery of foveal acuity.

Related experimentation has been concentrated in the following areas: (a) the investigation of the course of foveal dark-adaptation[1] and the effect of brief light-exposures on this light sensitivity;[2] (b) atomic-blast studies involving the effect of very brief (μ and m.sec.) and uniquely intense flashes on acuity;[3] (c) foveal acuity-studies after complete light-

* Received for publication April 3, 1964 .From the U.S. Naval Medical Research Laboratory, Groton, Connecticut.

[1] Selig Hecht, The nature of foveal dark adaptation, *J. gen. Physiol.*, 4 1921, 113-139; Selig Hecht, Charles Haig, and A. M. Chase, The influence of light adaptation on subsequent dark adaptation of the eye, *J. gen. Physiol.*, 20, 1937, 831-850.

[2] B. H. Crawford, Photochemical laws and visual phenomena, *Proc. Roy. Soc.*, Series B, 133, 1946, 63-75; F. A. Mote and A. J. Riopelle, The effect of varying the intensity and the duration of pre-exposure upon foveal dark adaptation in the human eye, *J. gen. Physiol.*, 34, 1951, 657-674; D. E. Johannsen, P. I. McBride, and J. W. Wulfeck, Studies on dark adaptation: I. The pre-exposure tolerance of the dark-adapted fovea, *J. opt. Soc. Amer.*, 46, 1956, 67-71; II. The pre-exposure tolerance of the human fovea adapted to different brightness levels, *ibid.*, 266-269.

[3] G. T. Chisum and J. H. Hill, Flash blindness recovery time following exposure to high-intensity short-duration flashes, U.S. Naval Air Development Center, Johnsville, Pa., 27 November 1961, NADC-MA-6142; S. L. Severin, Recovery of visual discrimination after high intensity flashes of light, School of Aerospace Medicine, USAF Aerospace Medical Center, Brooks AFB, Texas, December 1961, Report No. 62-16; S. L. Severin, N. L. Newton, and J. F. Culver, A study of photostress and flash blindness, USAF School of Aerospace Medicine, Brooks AFB, Tex., December 1962, SAM-TDR-62-144; J. F. Parker, Jr., Visual impairment from exposure to high intensity light sources, Office of Naval Research. Report prepared by Bio-Technology, Inc. under contract Nonr-4022(00). May 1963; S. L. Severin, A. V. Alder, N. L. Newton, and J. F. Culver, Photostress and flash blindness in aerospace operations, USAF School of Aerospace Medicine, Brooks AFB, Texas, September 1963, SAM-TDR-63-67.

adaptation of 5 to 10 min.;[4] and (d) the effect of brief light-exposures on peripheral light-sensitivity[5] and acuity.[6]

These studies indicate that foveal sensitivity to dim lights improves very rapidly with dark-adaptation, and that recovery of foveal dark-adaptation depends upon the intensity and the duration of the adapting source. They show, further, that the luminance required for resolution of a target, after complete light-adaptation, decreases with time in the dark. The curves are similar to those of simple foveal dark-adaptation, except that the final acuity-threshold is higher than the light-threshold and depends on the size of the target to be resolved. From these researches, we are also able to estimate the amount of time necessary to recover peripheral sensitivity after exposure to brief, bright lights.

Although these findings are important in themselves, they do not answer the question of how soon a dark-adapted fovea recovers its ability to discriminate targets in a darkened surround after exposure to a light flash of usual intensity and duration. It will be the purpose of this study to determine this function.

Apparatus. All measurements were made with a Hecht-Shlaer adaptometer.[7] The preadaptation-luminances were produced by the mechanism provided for this purpose in the instrument. A common source, in this case a 40-w. tungsten bulb, provided the light for both the preadaptation- and the test-stimuli. For the pre-

[4] J. L. Brown, C. H. Graham, Herschel Leibowitz, and H. B. Ranken, Luminance thresholds for the resolution of visual detail during dark adaptation, *J. opt. Soc. Amer.*, 43, 1953, 197-202; Brown, Effect of different preadapting luminances on the resolution of visual detail during dark adaptation, *ibid.*, 44, 1954, 48-55

[5] L. K. Allen and K. M. Dallenbach, The effect of light-flashes during the course of dark-adaptation, this JOURNAL, 51, 1938, 540-548; E. A. Suchman and H. P. Weld, The effect of light-flashes during the course of dark adaptation, this JOURNAL, 51, 1938, 717-726; Charles Haig, The course of rod dark adaptation as influenced by the intensity and duration of pre-adaptation to light, *J. gen. Physiol.*, 24, 1941, 735-751; George Wald and A. B. Clark, Visual adaptation and chemistry of the rods, *J. gen. Psychol.*, 21, 1937, 93-105; Mote and Riopelle, The effect of varying the intensity and the duration of pre-exposure upon subsequent dark adaptation in the human eye, *J. comp. physiol. Psychol.*, 46, 1953, 49-55; The effect of varying the light-dark ratio of intermittent pre-exposure upon subsequent dark adaptation in the human eye, *J. opt. Soc. Amer.*, 41, 1951, 120-121, S. M. Luria and J. A. S. Kinney, The interruption of dark adaptation, U.S. Naval Medical Research Laboratory, Groton, Conn., Vol. 20 (Report No. 347), 1 Feb. 1961.

[6] A. L. Diamond and A. S. Gilinsky, Luminance thresholds for the resolution of visual detail during dark adaptation following different durations of light adaptation, USAF Air Research Development Command, Wright-Patterson AFB, Aero Med. Lab. April 1952, WADC-Tech. Rep. 52-257; G. A. Fry and Mathew Alpern, Effect of flashes of light on night visual acuity. Wright Air Development Center, WADC Tech. Rep. 52-10, Part I, November 1951.

[7] Selig Hecht and Simon Schlaer, Adaptometer for measuring human dark adaptation, *J. opt. Soc. Amer.*, 28, 1938, 269-275.

adaptation-stimuli, the light from the source, projected by a series of lenses to a Maxwellian view, subtended approximately 35° at the O's position. Six settings of preadaptation-stimuli were used, ranging from 0.36 ft.-L to 3,000 ft.-L.

The test-stimulus was a circular acuity-grid, 1° in diameter, whose luminance could be varied by neutral filters and a neutral density-wedge. The bars of the grid, alternately opaque and transparent, subtended 6 minutes of visual angle yielding a visual acuity of 0.2. Two fixation-points were provided, one on either side of the acuity-grid, and O was instructed to fixate the center area between them where the test-stimulus would appear.

Procedure. The O first dark-adapted for 5 min.; then measures were made of his acuity-threshold by a method of constant stimuli. The adaptation-source was then presented for a given interval and the course of readaptation to the previously determined threshold was measured. Since foveal adaptation often proceeds very quickly, a single curve could not be determined all at once; therefore the following procedure was adopted: The acuity-grid was set at a predetermined level of luminance above final threshold and was presented repeatedly for one second at 5 sec. intervals until the O reported seeing it. If the first level was considerably above final threshold, a second lower luminance-level was then set and the procedure repeated. To fill in other points in the curve, the adaptation-source was presented again and the test-stimulus set at different luminance-values. Sufficient time was allowed between presentations of the adaptation-source to assure that the O always started from the completely dark-adapted condition. All the durations for a given adapting luminance were measured in one session with the order of presentation of durations randomized within the session. At least three sessions were run for each adaptation-level. Three Os, corrected when necessary to 20/20 vision, were run on all conditions. All observations were made with the right eye.

Results. The average threshold, after complete foveal dark-adaptation, for the 0.2 visual acuity-target was approximately 7.0 log $\mu\mu$L or 0.01 ft.-L. After exposure to the various adaptation-sources, the luminances required to resolve the target were increased, and varying lengths of time in the dark were needed to return to this thrshold. Fig. 1–3 show the recovery-curves averaged for the three Os for the various durations of constant luminance. For example, in Fig. 1, after 45 sec. of 3,000 ft.-L., the initial threshold was raised more than 2 log units and it required about 4 min. in the dark for the final threshold of 7.0 log $\mu\mu$L. to be regained. At the other extreme, after 1.5 sec. of 3,000 ft.-L., only 26 sec. were required for recovery.[8] The obvious differences between these families of curves for the various luminances of exposure are that the initial thresh-

[8] As long as the adapting stimulus fills the fovea, its size should not be a factor in foveal readaptation-time (J. A. Hanson, E. M. S. Anderson, and R. P. Winterberg, Studies on dark adaptation: V. Effect of various sizes of centrally fixated pre-exposure fields on foveal and peripheral dark adaptation, *J. opt. Soc. Amer.*, 50, 1960, 895-899). Nevertheless, as a check on this, the adapting field was cut to 4° for one unreported series of curves; the resulting data were the same as for the 35° field-size.

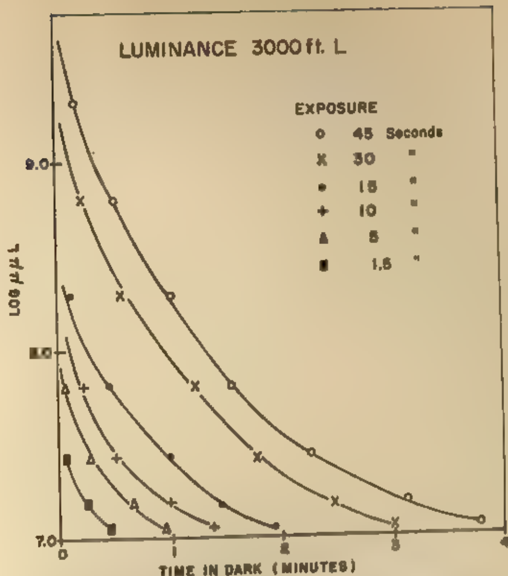

FIG. 1. READAPTATION FOLLOWING EXPOSURE TO 3,000 FT.-L. FOR VARIOUS DURATIONS

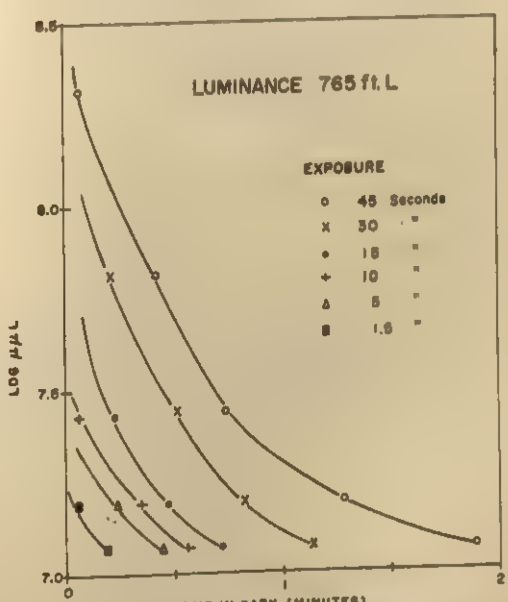

FIG. 2. READAPTATION FOLLOWING EXPOSURE TO 765 FT.-L. FOR VARIOUS DURATIONS

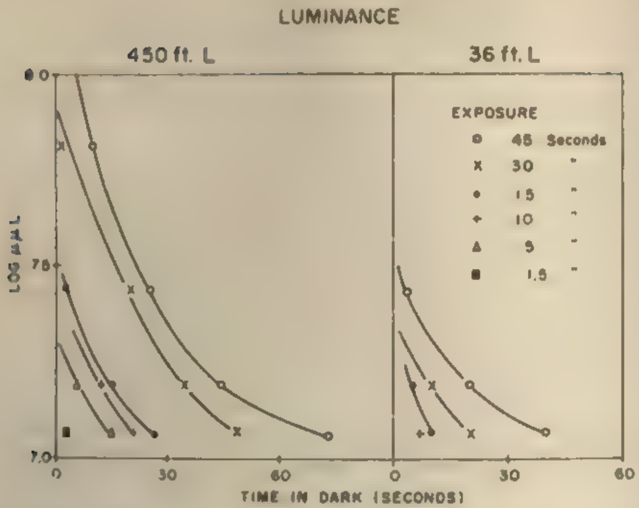

Fig. 3. Readaptation Following Exposure to 450 ft.-L. and to 36 ft.-L. for Various Durations

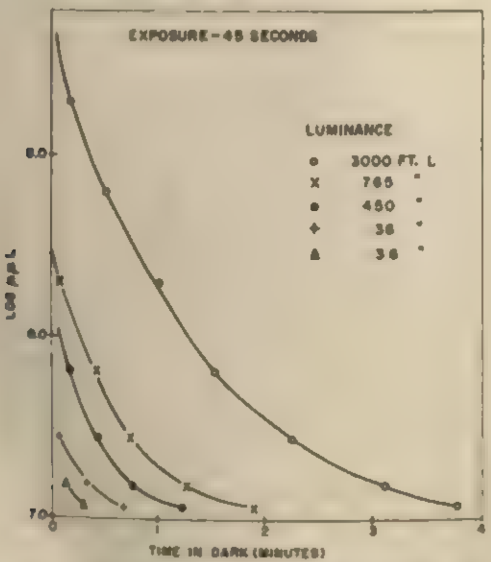

Fig. 4. Readaptation Following Exposure for 45 sec. of Various Luminances

old is higher and the total time to readapt longer for the brighter source.

This point can be made more explicit by plotting the same data with intensity the parameter rather than time. Fig. 4 is an example of this treatment for a duration of 15 sec. With this length of time it was possible to measure an effect of 3.6 ft L., but with 0.36 ft L., no effect on foveal acuity was found for any length of time.

The data presented thus far have been the averages of 3 Os and there are, of course, individual differences. Regular families of curves were

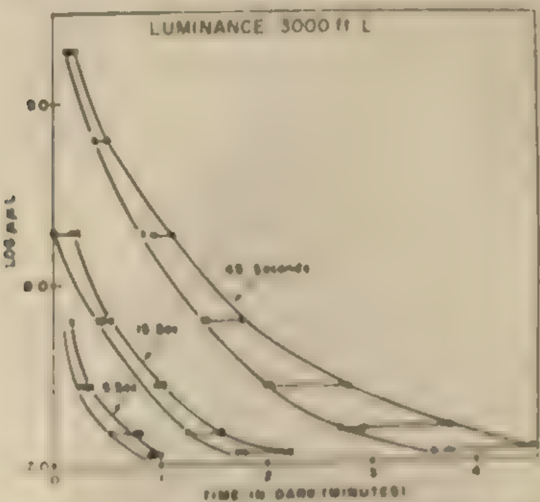

Fig. 5. Individual Differences in Readaptation Following Exposure to 3,000 ft L. for 5, 15, and 45 sec.

found, however, for each O: the initial threshold and the total time to readapt varied systematically with the intensity and duration of the adaptation source in the same way as was shown in the average curves. Fig. 5 is an example of the individual variations. Here the exposure times of 45, 15 and 5 sec. are given at 3,000 ft L., the luminance at which the largest individual differences were found. The horizontal lines indicate the range of these differences. Thus after 45 sec. the extreme values for the final threshold were 3.5 to 4.5 min. and after 5 sec. exposure, 30 to 60 sec. Individual differences are thus minor compared to the effects of the glare source.

Since the data suggested a reciprocal relationship between intensity and time of the adaptation source, the results for the various time intervals

combinations were tabulated and are shown in Fig. 6. The total time necessary for the fovea to readapt to threshold is plotted as a function of the log of the product of intensity and duration of the source. A single line was found to represent most of the data-points. The curve starts at a level of very little effect, 10 sec. or less of readaptation-time, and rises sharply to an asymptote of 4–5 min. with large values of intensity × time.

Discussion. Fig. 6 has shown reciprocity between intensity and time of exposure to light, as measured by the total time to readapt, for a re-

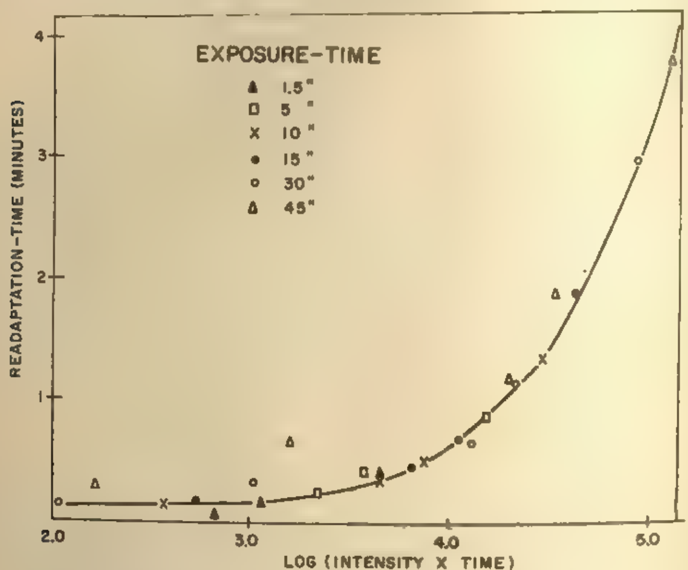

FIG. 6. RELATION BETWEEN READAPTATION-TIME AND LIGHT EXPOSURE AS MEASURED BY THE PRODUCT OF I × T

markable range of values. Constant products of I × T yield the same recovery-times for luminances varying at least 3 log units and for times of 1–30 sec. or more. While similar results have been shown for wide ranges of values for peripheral dark-adaptation,[9] the effect has not been reported previously for foveal dark-adaptation. Crawford states that complete

[9] Suchman and Weld, *op. cit.*, this JOURNAL, 51, 1938, 717-726; Mote and Riopelle, *op. cit.*, *J. comp. physiol, Psychol.*, 46, 1953, 49-55; Fry and Alpern, *op. cit.*, WADC Tech. Rep. 52-10, Part I, Nov. 1951; Crawford, *op. cit.*, *Proc. Roy. Soc.*, Series B, 133, 1946, 63-75; Haig, *op. cit.*, *J. gen. Physiol.*, 24, 1941, 735-751.

reciprocity holds for foveal vision only for durations of less than 1 sec., although failures in his data, in terms of total recovery-time, are obvious only at 30 sec. or longer.[10] Mote and Riopelle reported many instances of complete reciprocity for foveal dark-adaptation but also found a number of instances of failure.[11]

These studies differ from the present one in that they deal with foveal light-sensitivity rather than acuity. It is more important, perhaps, to note that there is a logical temporal limit beyond which the function cannot apply; that is, the exposure-time beyond which no further increase in recovery time is found.

Although this temporal limit will undoubtedly vary with the exposure-conditions, in the present experiment it was somewhat less than 1 min.; that is, recovery-times following 1-min. exposures were essentially the same as those following 45 sec. If a similar limit is assumed for the data of Mote and Riopelle, an assumption which is not at variance with their data,[12] the times required to reach a steady state, reported by them, give a curvilinear function against log I × T very similar to the one reported here.

The failure of reciprocity, where it occurs, is most interesting from a theoretical point of view. The only indication of failure in these data is at exposures of 45 sec., where recovery-times appear to be somewhat too long for simple reciprocity. This is in agreement with Crawford's data. He postulates a secondary photochemical process to account for the higher thresholds found with the longer exposures.[13]

From a practical point of view, however, the positive results are much more important. When a person is required to discriminate detail in a darkened environment, as in night driving, flying, or piloting a ship, his ability, after exposure to light, may be drastically reduced or may not be hindered at all. The amount of decrement depends upon the intensity and the duration of the exposure, and can be adequately predicted from the product of I × T.

Summary. The study was designed to measure the effect of various durations and intensities of light on the acuity-threshold of the dark-adapted fovea, by determining the time necessary to readapt following these exposures. The adapting lights were always presented foveally and

[10] Crawford, *op. cit.*, 69.
[11] Mote and Riopelle, *op. cit.*, *J. gen. Physiol.*, 34, 1951, 657-674.
[12] Mote and Riopelle, *op. cit.*, 1951, 665.
[13] Crawford, *op. cit.*, 75.

varied in brightness from 0.36 to 3,000 ft.-L., and in duration from 1 to 45 sec.

The resulting families of dark-adaptation curves show that the times necessary to readapt to the previously determined acuity-threshold vary systematically with the intensity and duration of exposure, from essentially zero for dim, brief lights to a maximum of about 5 min. for the longer, brighter ones. A most interesting aspect of the data is that the product of intensity and time gives a constant effect. When readaptation-time is plotted against I $\times$ T, a single curve results, which quite adequately fits the data-points.

THE EFFECTS OF ENDBOX-PLACEMENT ON SUBSEQUENT PERFORMANCE IN THE RUNWAY WITH COMPETING RESPONSES CONTROLLED

By R. C. GONZALEZ, Bryn Mawr College, and
BRYAN SHEPP, Brown University

In latent learning experiments of the placement variety, rats are trained in a runway leading to an empty endbox, then placed directly in the endbox (or in some similar box) with food, and then tested once more in the runway. Facilitation of running in the third stage of such an experiment by rewarded placement in the second stage may be predicted from rather different points of view. Cognitive theory suggests that facilitation will depend on the similarity of the feeding-box to the endbox.[1] Spence's revised theory of *S-R* reinforcement suggests that facilitation will depend on the similarity of the feeding-box to the runway.[2]

In a recent test of these divergent views by Gonzalez and Diamond,[3] one group of rats was trained in a runway (*R*) leading to an endbox (*B*) of the same color (*RxBx*) in Stage 1, while for another group the runway led to an endbox of a different color (*RxBy*). In Stage 2, half of each group was fed in *Bx* and the other half in *By*, yielding four experimental groups: I, *RxBx-Bx;* II, *RxBx-By;* III, *RxBy-Bx;* and IV, *RxBy-By.* To understand the performance of the animals in Stage 3, it was necessary to take into account *both* the similarity of feeding-box to endbox and the similarity of feeding-box to runway.

The first relation was exactly that which was predicted from cognitive theory: the greater the similarity of feeding-box to endbox, the greater the facilitation. (Group I ran faster than Group III, and Group IV ran faster than Group II.) The second relation was the opposite of that predicted from Spence's theory: the greater the similarity of feeding-box to runway, the less the facilitation. (Groups II and IV showed significant increases in speed of running, while Groups I and III did not; Group III, for which the feeding-box was the same color as the runway but a different color than the endbox, actually ran more slowly in Stage 3 than in Stage 1.) The latter results suggest that if the hypothetical r_G (assumed to be evoked in the runway by stimulus-generalization) plays any role in placement-experiments, the

* Received for publication March 1, 1964. This research was supported in part by Grant MH-9577 from the National Institute of Mental Health. The assistance of Peter Eimas in the conduct of Experiment II is acknowledged with gratitude.

[1] E. C. Tolman, Principles of purpose behavior, in Sigmund Koch (ed.), *Psychology: A Study of a Science,* 2, 1959, 95-112.

[2] K. W. Spence, *Behavior Theory and Conditioning,* 1956, 134-164.

[3] R. C. Gonzalez and Leonard Diamond, A test of Spence's theory of incentive-motivation, this JOURNAL, 73, 1960, 396-403.

role is a disruptive rather than a facilitative one. Since there is no reason to assume that r_G is incompatible with running, it may be that r_G serves a *reinforcing* function. The tendency of animals to loiter in a runway of the same color as a box in which they have been fed certainly seems understandable on that assumption.

Experiment I

However the disruptive effect of feeding-box-to-runway similarity is to be interpreted, the facilitative effect of feeding-box-to-endbox similarity found by Gonzalez and Diamond has considerable importance for the old question of whether the rat learns *about* the consequences of its responses. The experiment now to be reported was designed to examine the facilitative effect in a purer form, uncomplicated by the disruptive effect. Three groups of rats were studied for which the similarity of feeding-box to runway was held constant while the similarity of feeding-box to endbox was varied systematically.

Subjects. The *Ss* were 24 experimentally naïve hooded rats, 130-150 days old, obtained from the Department of Zoology at the University of Connecticut.

Apparatus. The apparatus consisted of a starting-box (6 in. long × 6 in. wide), a runway (68 × 4 in.), and three interchangeable endboxes (12 × 6 in.). The endbox in use on any given trial was fixed at a 90° angle to the end of the runway and entered by a right turn. (The right-angle made the interior of the endbox invisible from the runway.) Each unit was made of wood, with walls 6 in. high, and had a hinged Plexiglas cover. The walls and floors of the starting-box, the runway, and one of the endboxes were painted flat mid-gray. The walls and the floor of another endbox were painted flat black, and those of the third endbox were painted flat white. A manually operated sliding door, which served to prevent retracing, was set into the runway 3.5 in. before the endbox. Recessed in one wall of the runway, 1 in. beyond the starting-compartment, was a photocell upon which fell a light-beam from the opposite wall. When this beam was broken by the animal, two Standard Electric timers (graduated in units of 0.01 sec.) were started. The first timer was stopped when the animal broke a second light-beam which crossed the runway 19.5 in. beyond the first, and the second timer was stopped when the animal broke a third beam which crossed the runway 2.5 in. before the goal-compartment. Two measures of running-time thus were afforded—*initial time* and *total time*.

Preliminary training. For a few days after their arrival in the laboratory, the animals were maintained on a free-feeding schedule, and their satiated bodily weights were determined. Then they were placed on a 24-hr. feeding schedule, and their weights were reduced gradually to the 80% level. The animals were taken once daily, for 20 days, into the experimental room for a 15-min. period of handling; 2 hr. later, each *S* was fed in its home-cage the amount of Purina chow required to maintain its weight at the 80% level. On the basis of their adjustment to handling, the animals were divided into three matched groups of 8 *Ss* each: Group *W*, Group *G*, and Group *B*.

Stage 1. In the first stage of the experiment-proper, there were 20 trials—all

unrewarded one on each of 20 successive days. *S* was taken to the experimental room and allowed to adapt for 2 min. in the carrying cage before being placed in the apparatus. At the start of each trial, the door used to prevent retracing was open; *S* was placed in the starting-box and, as soon as it broke the third light-beam, the door was closed. *S* remained in the endbox for 1 min., after which it was returned to its home-cage; its daily ration was fed 2 hr. later. Group *W* found the white box at the end of the runway, Group *G* the gray box, and Group *B* the black box. On the basis of their performance during this stage of training, the animals of each group were divided into two matched subgroups of 4 *S*s each, one of which was to receive its placement-feedings in the black box and the other in the white box.

Stage 2. During the second stage of the experiment, which lasted for five days, *S* was placed once each day in one of the endboxes (either black or white) and fed. After being brought into the experimental room, *S* was allowed 2 min. in

TABLE I

DESIGN OF EXPERIMENT I

Group	*N*	Runway	Endbox	Feeding-box
I (Same)	4	gray	black	black
	4	gray	white	white
II (Intermediate)	4	gray	gray	black
	4	gray	gray	white
III (Different)	4	gray	white	black
	4	gray	black	white

which to adapt in the carrying cage, and then placed directly in the endbox which contained a small cup with 6 oz. of wet mash. A partition, painted the same color as the endbox, prevented entrance into the alley. After 4 min., *S* was removed from the endbox, returned to its home-cage, and, on the first four days, fed the remainder of its daily ration 2 hr. later. On the fifth day, the first of the series of Stage-3 test-trials preceded the supplementary feeding.

Group I (Same) consisted of four animals of Group *W* fed in the white box and four animals of Group *B* fed in the black box. Group III (Different) consisted of four animals of Group *W* fed in the black box and four animals of Group *B* fed in the white box. Group II (Intermediate) consisted of the eight animals of Group *G*, half of which were fed in the white box and half in the black box. The experimental conditions for each group are summarized in Table I.

Stage 3. On Day 25 of the experiment, 2 hr. after its fifth placement-feeding, *S* was returned to the experimental room, allowed 2 min. in which to adapt, and then placed in the starting-box for the first test-trial. The apparatus and procedure for each *S* were exactly the same as in the first stage of training. In all, there were eight test-trials, one on each of eight successive days, all unrewarded.

Results. The performance of the three groups in Stages 1 and 3 is shown in Fig. 1 in terms of mean log initial time on each trial. The plots for mean log total time, very much the same, are not shown in the interest of

conserving space. In the first stage, the three groups did not differ significantly in terms of either measure ($F < 1$). The effects of the placement-feedings on subsequent running were evaluated separately for each group by means of t-tests of the difference in time between the last trial of Stage 1 and the first trial of Stage 3. As measured in terms of initial time, both Group I ($t = 5.60$, $p < 0.01$) and Group II ($t = 5.04$, $p < 0.01$) showed significant decreases; the decrease for Group III approached but did not reach significance ($t = 2.18$, $0.10 > p > 0.05$). The magnitude

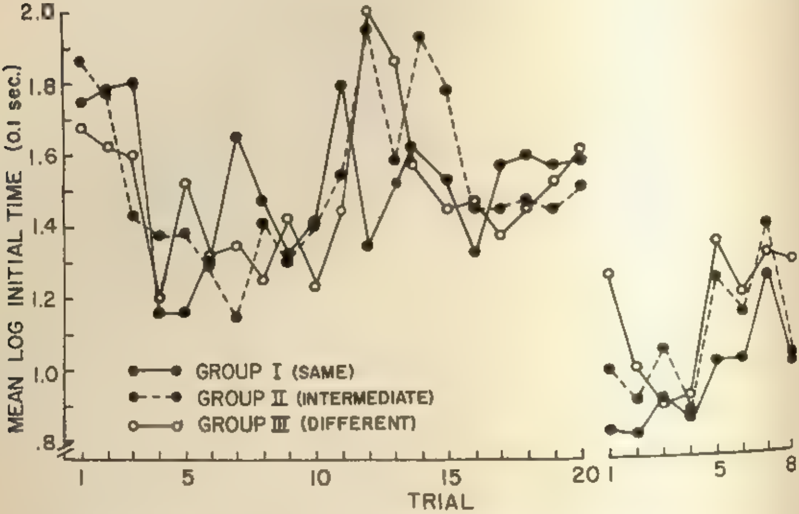

FIG. 1. THE PERFORMANCE OF THE THREE GROUPS IN STAGES 1 AND 3 OF EXPERIMENT I MEASURED IN TERMS OF INITIAL TIME.

of the decrease, furthermore, was significantly greater for Group I than for Group II (for the difference between the difference-scores of the groups $t = 2.58$, $p < 0.05$). Measured in terms of total time, only the decrease for Group I was significant ($t = 5.13$, $p < 0.01$). The decrease for Group II approached significance ($t = 2.03$, $0.10 > p > 0.05$), while that for Group III was negligible.

In this experiment, with the same runway-stimuli, the same feeding-box stimuli, and the same extra-maze stimuli for all animals, speed of running in Stage 3 varied directly with the similarity of the empty endbox to the feeding-box. Clearly, then, the instrumental response may be activated differentially quite independently of the tendency for the runway to evoke r_G. The results may be understood on the assumption that the animals

learned in the first stage that the runway led to an endbox of a particular color, and that the feedings in the second stage led to the anticipation of food in that box, the strength of the anticipation varying directly with its resemblance to the feeding-box.

It is possible to deal with these results in *S-R* terms if some rather cumbersome assumptions are made about the chaining of implicit responses. To account for learning about the properties of the endbox, it might be assumed that the box evokes some characteristic (perceptual) response (r_{eb}) which becomes connected to the proprioceptive feedback from the instrumental response (s_R) with which it is continguous, this connection developing in Stage 1 concurrently with the connection of the instrumental response (R) to the runway-stimuli (S). To account for the facilitative effect of feeding-box-to-endbox similarity, it might be assumed that the feeding-box evokes a characteristic response of its own (r_{fb}), whose proprioceptive consequences become connected during placement to the ensuing consummatory response. The sequence of events in Stage 3 then would be characterized as follows: $S \rightarrow R — s_R \rightarrow r_{eb} — s_{r_{eb}} \rightarrow r_G — s_G$, the instrumental response being facilitated to the extent that it produces r_G. An important link in the chain leading from R to r_G is the generalized connection $r_{eb} \rightarrow r_G$, whose strength, is a function of the similarity of feeding box to endbox. (To the extent that the feeding-box and the *runway* are similar, the runway will evoke r_G and interfere with instrumental response.)

It is possible in principle to account for the effects of placement in these terms, however contrived and unparsimonius, as it is not in terms of r_G as a generalized activator of instrumental behavior. (The account need not, of course, be tied to a consummatory or a need-reduction interpretation of reward; a positive affective response would do as well.) Whether or not an *S-R* interpretation of this sort has anything to recommend it over an *S-S* interpretation remains to be determined. In the meantime, it is significant that the results of these placement-experiments follow simply and directly from cognitive theory.

Experiment II

In Experiment II, the generality of the results obtained in Experiment I was evaluated in a latent-extinction design.[4] In Stage 1, running was *rewarded;* in Stage 2, the animals were placed in *empty* endboxes; in Stage 3, the effects of placement were sought in *decreased* speed of running. Two groups were studied for which, as in Experiment I, the similarity

[4] J. P. Seward and Nissim Levy, Sign learning as a factor in extinction, *J. exp. Psychol.*, 39, 1949, 660-668.

of the placement-box to the runway was held constant while its similarity to the endbox was varied.

Subjects. The *Ss* were 24 experimentally naïve hooded rats, 90-120 days old.

Apparatus. The apparatus was that of Experiment I, with the exception that only two endboxes—black and white—were used.

Preliminary training. After preliminary handling, and the reduction of their weights to the 85% level, the animals were divided into two groups of 12 *Ss* each —Group *W* and Group *B*—matched for adjustment to handling and weight. Group

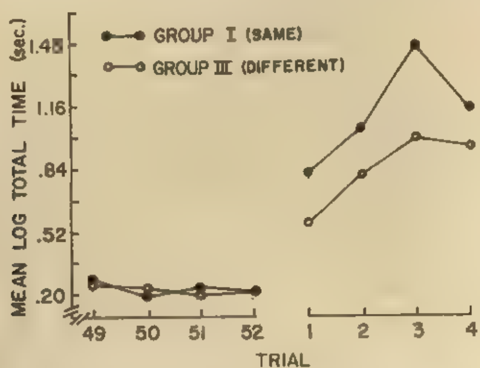

FIG. 2. THE PERFORMANCE OF THE TWO GROUPS ON THE LAST FOUR TRIALS OF STAGE 1 AND THE FOUR TRIALS OF STAGE 3 IN EXPERIMENT II.

W found the white box and Group *B* the black box in the first stage of experimental training.

Stage 1. In this stage there were 52 rewarded training trials, four on each of 13 consecutive days. On each trial, *S* was permitted to eat for 30 sec. from a cup of wet mash in the endbox, after which it was placed in a detention-cage for 1 min., and then placed in the starting-box for the beginning of the next trial. After the last trial of each day, *S* was returned to its home-cage and fed the remainder of its daily ration 2 hr. later. On the basis of performance during this stage of training, the animals of each group were divided into two matched subgroups of 6 *Ss* each, one of which was to receive its placements in the empty white box, and the other in the empty black box.

Stage 2. On Day 14 of the experiment, each *S* was placed in one of the endboxes—now empty—for four 3-min. periods; the interval between placements was 1 min. A partition, painted the same color as the endbox, prevented entrance into the runway. Following the fourth placement, *S* was returned to its home-cage, where it was fed 2 hr. later. On Day 15, each *S* was given an additional placement in the empty endbox. Group I (Same) consisted of six animals of Group *W* placed in the empty white box and six animals of Group *B* placed in the empty black box; Group III (Different) consisted of the remaining 12 animals.

Stage 3. Immediately after the fifth placement, *S* was put in the detention-cage;

1 min. later, *S* was placed in the starting-box for the first test-trial. There were four test-trials, all given on Day 15. The apparatus and procedure were the same as in Stage 1, except that none of the test-trials was rewarded.

Results. Because some unreliability developed in the second photocell during the test-trials, only the data for total time are presented. Mean log total time for the two groups is shown in Fig. 2 for the last four trials of Stage 1 and the four test-trials of Stage 3. Both groups showed significant increases in running time (for the difference between time on the last trial of Stage 1 and the first trial of Stage 3, $t = 6.22$, $p < 0.01$, for Group I, and $t = 4.37$, $p < 0.01$, for Group III). The increase was significantly greater for Group I than for Group II ($t = 2.11$, $p < 0.05$).

In this experiment, then, the magnitude of the effect of placement on subsequent performance varied directly with the similarity of the endboxes encountered in the first two stages of training—just as it did in Experiment I. This outcome may be understood in terms of the same set of principles from cognitive theory, unrewarded placement leading to a reduction in the valence of the endbox of Stage 1 in proportion to its similarity to the placement-box. Taken alone, of course, the results of this experiment are not crucial to Spence's r_G-interpretation. It can be assumed that for the animals of Group I the *rate* of emission of r_G in Stage 2 was greater, the amount of extinction of r_G was greater, and, therefore, the value of K in the third stage was less. With the results of Gonzalez and Diamond and those of Experiment I showing clearly that K does not activate instrumental response in the manner suggested by Spence, it seems pointless to consider such an explanation here.

SUMMARY

Two placement-experiments are reported, one of a latent-learning and the other of a latent-extinction design. In both experiments, with the similarity of the placement-box to the runway constant for all animals, the magnitude of the effects of placement on subsequent performance were found to vary directly with the similarity of the placement-box to the endbox encountered in the first stage of training. The implications of the results for *S-S* and *S-R* theories of learning are considered.

THE RELATION BETWEEN PULSE-TO-CYCLE FRACTION AND *CFF*

By C. I. Thompson, F. G. Fidura, C. J. Rhoads, C. A. DeYoung, and B. L. Kintz, Ohio University

The classical notion about the role of the pulse-to-cycle fraction (*PCF*) of an intermittent photic pulse is that, once *PCF* is increased enough to eliminate flicker, further increases will result only in increased brightness and never in the reappearance of flicker. Recent evidence casts doubt on the validity of this notion.

Bartley and Nelson found that *PCF*-curves cross when *CFF* is plotted against intensity.[1] From this fact they inferred that more than one *PCF* could produce the transition between flicker and fusion at a given *CFF* and intensity. Using Hartline's findings with respect to the off-response in the optic nerve,[2] Bartley postulated a neurophysiological model in which changes in *PCF* would produce three transition-points.[3] Specifically, the model provided that at a given frequency, using an intermediate intensity, low *PCF*-values would produce flicker, which then would change to fusion as *PCF* was increased. Still further increases of the *PCF* would reintroduce flicker, and, at the highest *PCF*-values, a final fusion would appear.

An attempt to observe the three transitions directly was made by Bartley and Nelson, who used an electronic apparatus to vary *PCF* while holding frequency and intensity constant.[4] These investigators found that their procedure produced more than one transitional point, but at intermediate intensities there were only two transitions rather than the expected three. Low *PCF*s produced fusion, which changed to flicker, and then again to fusion as *PCF*-values were increased. On the basis of these results, Bartley and Nelson speculated that only with very high intensities could short pulses produce flicker, although Bartley's earlier formulation had suggested that only one transition was to be expected at high intensity.

The disagreement of these results with the proposed model suggests the need for a further investigation. The present study was specifically designed to test the prediction of three transition-points.

Observers. The first four authors, two men (*Tho* and *Fid*) and two women

* Received for publication January 11, 1965.

[1] S. H. Bartley and T. M. Nelson, Some relations between pulse-to-cycle fraction and critical flicker frequency, *Percept. mot. Skills*, 10, 1960, 107-115.

[2] H. K. Hartline, The response of single optic nerve fibers of the vertebrate eye to illumination of the retina, *Amer. J. Physiol.*, 121, 1938, 400-415.

[3] S. H. Bartley, Some factors influencing critical flicker frequency, *J. Psychol.*, 46, 1958, 107-115.

[4] Bartley and Nelson, A further study of pulse-to-cycle fraction and critical flicker frequency: A decisive theoretical test, *J. opt. Soc. Amer.*, 51, 1961, 41-45.

(*Dey* and *Rho*), served as *O*s in this study. They had previous experience in at least one *CFF*-study prior to the present one.

Apparatus. Photic pulses were presented by a Sylvania glow-modulator tube which was driven by a Grayson-Stadler fusion apparatus (Model E622). With this apparatus it was possible to vary *PCF* from 0.02-0.98 and at the same time keep flicker-frequency and intensity constant at any desired level. The light-source was 3/32 in. in diameter, and subtended a visual angle of 1°, centered on *O*'s left fovea. It was necessary to keep *O*'s head fixated in one position so that the entire presentation could be utilized in making the discrimination. This was accomplished by mounting a wooden viewer, similar to that of a tachistoscope, in such a position that *O* was always fixated on the point of light when looking into the viewer. A rubber unit molded to fit the contours of the face surrounded the wooden viewer in such a manner that *O* could lean his forehead against it, and comfortably maintain a viewing position over long periods of time. The right half of the unit was blocked to permit stimulus-presentation to only the left eye. The *O*s were permitted to wear corrective lenses if they normally did so. All trials were given in a dark room, and further protection from extraneous sources of light was provided by enclosing *O* in a cubicle which was painted flat black.

Procedure: Part I. The experiment was conducted in two parts, the first of which consisted of an attempt to locate points at which *PCF*-curves crossed when *CFF* was plotted against intensity. *PCF*s of 0.10, 0.30, and 0.50 were paired with intensities of 11.3, 16.0, and 22.6 ma., and *CFF*s were found for each of the nine resulting combinations. The method of limits was used to determine *CFF*s in each combination, with the average transition-point being taken from two ascending and two descending series. Before beginning the experiment, all *O*s were dark-adapted for a minimum of 5 min. The burst-duration was 1 sec., and intertrial intervals were 5 sec. All measures for each *O* were taken during a single session.

Results: Part I. Fig. 1 shows the typical results obtained and will serve to illustrate the findings for all three *O*s who showed *PCF*-crossings. For the particular *O* shown in Fig. 1, it will be noted that the curves for the *PCF*s of 0.50 and 0.30 crossed at a frequency of approximately 38.5 c.p.s. and an intensity of 16 ma. The curves for the *PCF*s of 0.50 and 0.10 crossed at approximately 36 c.p.s. and 16 ma. Although the curves crossed at other points for the other two *O*s, the same basic procedure illustrated above was used in determining the flicker-frequencies and intensity-levels to be held constant during the second part of the study.

For all *O*s, the variability of the transitional points observed was very low; the largest difference within any set of four series being 4 c.p.s. These findings are typical of those reported for trained *O*s when *CFF* is determined in this manner.

Procedure: Part II. The *CFF*s and intensity-values of the crossing points were held constant, and *PCF* was varied from 0.02-0.98 in the method of limits. For the one *O*, for whom no crossing of the *PCF*-curves was obtained in Part I, the experiment was continued at an intermediate *CFF* and intensity, which was held constant. The

intertrial interval again was 5 sec. To allow the *O*s enough time to make the discrimination, however, it was necessary to increase the burst-duration to 1.5 sec.

Results: Part II. Table I indicates the intensities and flicker-frequencies which were held constant while *PCF* was varied, the proportions of series for which low *PCF*-transitions occurred, the mean *PCF* at the transitions, and the *SD*s. It is interesting to note that results for the *O* who did not show crossing in Part I (*Fid*) are very similar to results obtained for the other *O*s, which may mean that the preliminary phase of this study was unnecessary. The table also gives like data for the middle and high *PCF*-transitions.

TABLE I

MEANS AND *SD*s OF TRANSITIONAL POINTS FOR EACH *O*
(Numbers in parentheses refer to the proportion of series for which transition occurred.)

O	Intensity and flicker-frequency	Fusion Mean	Fusion *SD*	Flicker Mean	Flicker *SD*	Fusion Mean	Fusion *SD*
Tho	19 ma., 40.4/c.p.s.	0.063 (6/8)	0.019	0.313 (6/8)	0.076	0.550 (6/8)	0.087
	22.6 ma., 40.4/c.p.s.	—	—	.143 (3/4)	.068	.712 (4/4)	.043
Dey	16 ma., 38.6/c.p.s.	—	—	.289 (8/8)	.095	.528 (8/8)	.078
Fid	22.6 ma., 38/c.p.s.	0.040 (1/4)	0.000	0.172 (4/4)	0.074	0.556 (4/4)	0.059
	22.6 ma.,/40 c.p.s.	—	—	0.240 (2/2)	0.028	0.550 (2/2)	0.071
Rho	22.6 ma., 36/c.p.s.	0.068 (4/6)	0.018	0.190 (6/6)	0.058	0.670 (6/6)	0.096
	16 ma., 38.5/c.p.s.	—	—	0.220 (3/4)	0.035	0.558 (3/4)	0.072

In all, 9 of the series showed three transitions, 23 showed two transitions, 3 showed one transition, and in one series no transitions occurred. For all series which showed three transitions, low *PCF*-values produced flicker, which progressively changed to fusion, back to flicker, and then to a final fusion as *PCF* was increased. For the series in which two transitions occurred, low *PCF*-values produced fusion, which changed to flicker and then back to fusion as *PCF* was increased. Two of the one-transition series resulted in a change from flicker to fusion at low *PCF*-values (0.08 and 0.06). These results are included in Table I with the first-transition data of the three-transition series. The other one-transition series changed from flicker at a low *PCF*-value, to fusion at a high *PCF*-value (0.65). This

datum is included in Table I with the final-transition data. In all series, the highest values of *PCF* produced fusion, which is the result to be expected as light occupies an increasing proportion of each cycle.

The means and *SD*s for the transitions were calculated only for those series in which transitions occurred. It is immediately obvious that transitions observed in this manner are not the stable measures obtained in conventional *CFF*-studies, where *PCF* is held constant. The variability would be even greater had those series in which no transitions occurred been combined with these measures.

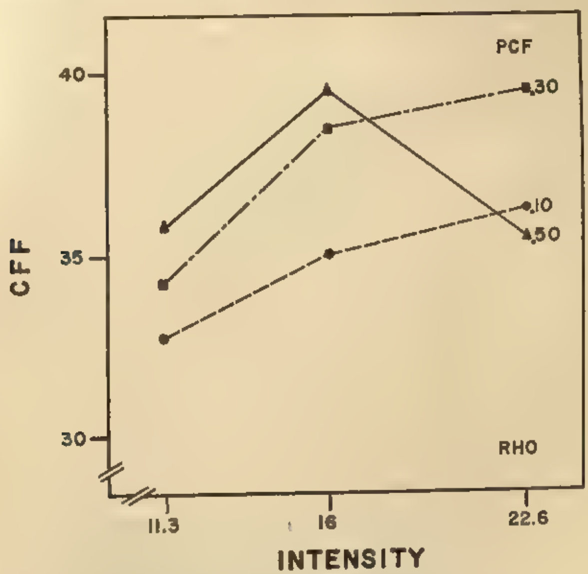

FIG. 1. CROSSOVER DATA FOR OBSERVER *Rho*
The intensity values are expressed in ma. of current passing through the stimulus-light.

Smoothed curves illustrating both two-transition and three-transition data are presented in Fig. 2. The horizontal line in each section represents the area above which flicker appeared and below which *O*s reported fusion. Points plotted in the top section of the figure are based either upon four or six transitions which occurred during a total of eight series. For the *O* illustrated in the bottom section, all eight series produced two transitions. Points in both sections which do not lie on the baseline are hypothetical and merely show the direction of the transition. The lines showing the final fusion in both sections of Fig. 2 could of course be extended to a *PCF* of 0.98.

Discussion. The results of this study clearly support Bartley and Nelson's conclusion that more than one transition are observable when *PCF* is varied while *CFF* and intensity are held constant.[5] Variable as the measures were, in only four of the 36 series did fewer than two transitions occur. A probable explanation for the four discrepancies is that the discrimination was so difficult that in these cases it was not made at all.

Even more interesting, however, is the fact that three *Os* reported three

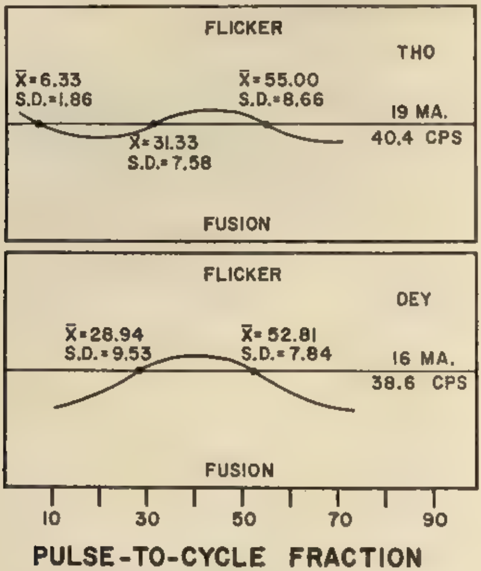

FIG. 2. TWO-TRANSITION AND THREE-TRANSITION DATA FOR OBSERVERS *Tho* AND *Dey*
All *PCF*-values, means, and *SDs* have been multiplied by 100.

transitions, a result not obtained by Bartley and Nelson. For the two *Os* who reported more than one series of three-transitions, *SDs* for the first transition were quite small, having values of 0.019 and 0.018. This low variability, coupled with the large proportion of transitions occurring within the total series, makes it unlikely that these results may be attributed to chance.

Bartley proposed that flicker will occur at low *PCFs* due to the fact that a continuous sensory result cannot be maintained when impulses of a very

[5] Bartley and Nelson, *op. cit., J. opt. Soc. Amer.*, 51, 1961, 41-45.

brief nature are separated by relatively long periods of darkness.[6] Further increases of *PCF* produce a summative effect on the retina, which causes the stimulus to appear fused. After reaching this point, further increase in *PCF* increases the sensory brightness of the stimulus, until the intensity becomes high enough to elicit an off-response in the optic nerve during the brief 'off' phases. Finally, the higher values of *PCF* allow such short periods of darkness between succeeding stimuli that the off-response is inhibited by the next flash of light and a final fusion is reached. This model, which easily handles three transition-points, provides the most acceptable explanation of the phenomenon.

There remains only the question of why fewer than three transitions occurred in many of the series. One possible answer was suggested by Bartley and Nelson, who noted that low and intermediate intensities show an inverted-*U* function when *CFF* is plotted against *PCF*.[7] They concluded that since fusion for low *PCF* values occurs at such a low flicker-rate, it is extremely difficult to produce a flicker-experience at those low *PCF*-values. The prediction which follows from this line of reasoning is that increasing the intensity, which gradually smoothes out the inverted-*U*, would produce conditions more favorable to flicker at low *PCF*-values. The results for *Tho,* shown in Table I, are not in line with this prediction. In both sets of data, flicker-frequency was held constant at 40.4 c.p.s. In the set which included the three-transition data, however, the intensity actually was decreased from 22.6-19.0 ma. From the preceding argument, it is expected that the three transitions would be most likely to occur at the higher intensity.

An alternative explanation of fewer than three transitions, and one which also has been given by Bartley and Nelson, is that a *PCF* of 0.02 is just not low enough to produce flicker. Although it will be noticed that for the three *O*s showing the three transitions at low *PCF*-values flicker was obtained with *PCF*s as high as 0.09, it way well be that *PCF*s lower than 0.02 would produce an even greater number of transitions than were obtained in the present experiment.[8]

---

[6] Bartley, *op. cit.,* 107-115.
[7] Bartley and Nelson, *J. opt. Soc. Amer.,* 51, *op. cit.,* 1961, 41-45.
[8] A fourth transition to no flicker (fusion?) should, however, occur at very low *PCF*-values, because the stimulus at a given flash rate and a given intensity will not emit enough photons in one flash to activate the receptor. At least, there will be no more than a temporally summated, infrequent visual experience. During some exploratory work, we noted exactly that phenomenon when (a) the intensity was low, (b) the flash-rate was low, and (c) the *PCF* was low. By increasing the intensity or the flash-rate or the *PCF,* we could increase the frequency of the summated visual experiences until a regularly flickering light was perceived.

Summary. PCF-crossings were obtained by varying flicker-frequency for nine combinations of *PCF* and intensity. After the crossing points were determined, the corresponding *CFF*- and intensity-values were held constant while *PCF* was varied from 0.02-0.98. The majority of these series produced two flicker-to-fusion or fusion-to-flicker transitions, while three transitions occurred for three *O*s. Bartley's original neurophysiological model provides a fairly accurate explanation of the retinal behavior involved.

DRIVE-LEVEL AND STIMULUS-IMPOVERISHMENT

By Herman Feldman, Indiana University, Northwest Campus

In recent years, attention has been focused on a possible relationship between motivational variables and perception. The results of these investigations have been interpreted in various ways due to differences in the definition of perception which make it difficult to compare the observations of many investigators. Perception is defined operationally in the present study as a recognitive response to stimuli which are tachistoscopically presented. The success of hypotheses concerning recognitive responses elicited under optimal stimulus-conditions has been summarized by Spence.[1] The study reported here is an attempt to extend Hull-Spence multiplicative drive theory to responses elicited by impoverished stimuli. Brown has discussed the relationship between stimulus-impoverishment and performance under high and low levels of drive and formulated the basis for a prediction which is tested in the study described below.[2]

Assuming that "impoverishment produces no change other than a decline in habit strength," Brown predicts that performance should improve under increased drive.[3] In addition, however, the role of competing responses may be important, since impoverished stimulation of the correct response might result in a rise in the strength of incorrect responses under conditions of high drive. This would have the effect of impairing performance compared to the same conditions under low drive.

This alternate prediction is the major hypothesis tested in this study. In this investigation, stimulus-impoverishment is accomplished tachistoscopically; thus, the greater the habit-strength of the response required in the tachistoscopic recognitive task, the lower will be the stimulus-intensity necessary to evoke that response, with drive-level held constant.

Drive-level is here defined in terms of high and low scores on the Taylor Anxiety Scale (*TAS*).[4] Since, in a task which involves the evoca-

---

* Received for publication August 17, 1964. This report is based upon a portion of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Nebraska. The writer wishes to express his appreciation to Professors Katherine Baker and Frank Dudek for their assistance.

[1] K. W. Spence, *Behavior Theory and Conditioning*, 1956, 221-235.

[2] J. S. Brown, *The Motivation of Behavior*, 1961, 271-280.

[3] Brown, *op. cit.*, 281-282.

[4] J. A. Taylor and K. W. Spence, Relationship of anxiety level to performance in serial learning, *J. exp. Psychol.*, 44, 1952, 61-62.

tion of competing response-tendencies, the probability that the correct response will occur on the first trial decreases as drive increases, recognitive thresholds will be higher for high-drive than for low-drive *Ss*.

From these theoretical considerations, the following hypotheses were generated and tested:

Hypothesis 1. Values of visual recognitive threshold for verbal material are a negatively accelerated function of the frequency with which visually presented verbal material has been reinforced, with drive-level held constant.

Hypothesis 2. Visual recognitive threshold values for verbal material will vary directly with drive-level with habit-strength held constant.

Method: Subjects. Drive-level was manipulated by selecting *Ss* on the basis of their scores on the *TAS.* The scale was given to 516 students in introductory psychology and sociology at the University of Nebraska. That intelligence was not an influential variable was evidenced by the fact that the mean *ACE* scores for high-anxious and low-anxious *Ss* were not significantly different. Students who had "Lie-Scale" scores of 6 or above were eliminated from consideration. The remaining 430 had scores ranging from 0 to 39, with a mean of 14.84. From this group, 35 high-anxious (*HA*) and 35 low-anxious (*LA*) *Ss* were chosen. Scores for the *HA* group ranged from 25 to 39, scores for the *LA* group from 0 to 6. These two groups represented approximately the upper and lower 16% of the total distribution.

Acquisition-Procedure. This phase of the study was introduced to the *Ss* as an experiment concerning foreign language training. Each *S* was presented with pronounceable seven-letter nonsense-words, taken from Solomon and Postman,[5] each word typed on a separate index card. The *S* was instructed to pronounce, one at a time, the 10 core words and 14 other filler words as they appeared randomly in a deck of 100 shuffled cards. The core words appeared or reappeared 1, 2, 5, 10, and 25 times. Two core words were used at each frequency. After each pronunciation, *E* said either "correct," "that's right," or "O.K." For each *S*, the frequencies of the latter words were systematically varied to counterbalance intrinsic differences in recognizability and pronounceability.

Threshold-Measurement. The frequency-category (habit-strength) of zero was composed of the average threshold of three nonsense-words which *S* had never seen during the acquisition-trials which preceded the measurement of the thresholds.

For the measurements, *S* was seated 8 ft. from a 35 × 30-in. sandblasted glass screen in a completely dark room. The test-words were projected on to the back of the screen from an adjoining room. *S* was instructed to keep his eyes directed at the screen, especially following each ready signal, and to report through an intercommunication system between the two rooms after each flash on the screen. A response of 'No' was to be given if the word was not recognized; otherwise *S*

[5] R. L. Solomon and Leo Postman, Frequency of usage as a determinant of recognition thresholds for words, *J. exp. Psychol.,* 43, 1952, 195-201.

was to say what he thought the word was. The threshold-test of Solomon and Postman was followed with respect to practice-words, order of presentation, and the inclusion of 10 English and 10 extra nonsense-words among the core words.

The projection consisted of a modified Viewlex slide-projector situated 6 ft. from the glass screen and directly in line with the *S* sitting in the adjoining room. Stimulus-words, on 2 × 2-in. transparent slides, were so focused on the screen that the typewritten letters measured 1⅛ × 1½ in. To obtain the proper range of luminance, the voltage across the 150-w. lamp of the projector was reduced, and a Corning glass daylight-filter (No. 5900) was mounted immediately in front of the lens. A piece of ¼-in. milk glass placed just in front of the projector-lamp served to diffuse the light evenly over the screen. The projector was fitted with a No. 4 Ilex Universal timing shutter set for an exposure of 0.2 sec. on all trials.

Control of luminance-level was accomplished by means of polaroid disks contained in a special mounting placed immediately in front of the lens and daylight

TABLE I

THRESHOLD-LUMINANCE (FT.-L.) FOR RECOGNITION OF NONSENSE-WORDS WITH DIFFERENT FREQUENCIES OF PRIOR EXPOSURE AT TWO DRIVE-LEVELS

f	Mean Log-L		SD Log-L		Geom. mean-L	
	high	low	high	low	high	low
0	− .954	− .979	.081	.062	.111	.105
1	−1.006	−1.030	.051	.048	.099	.093
2	−1.008	−1.040	.061	.054	.098	.091
5	−1.040	−1.074	.081	.049	.091	.084
10	−1.038	−1.083	.078	.050	.092	.083
25	−1.065	−1.116	.067	.055	.086	.077

filter. The angular displacement of the axes of the rotating- and the stationary-disk could be read to the nearest 1° from a calibrated scale. The maximal luminance was measured to be 0.1425 ft.-L with a Macbeth Illuminometer placed on *S*'s side of the screen. Other levels of luminance less than the maximum were found from the relationship between the angle of displacement of polaroids and measured luminance obtained at 5° intervals over the scale. Care was taken to avoid very low settings where color-changes are likely to occur.

For the threshold-determinations, the polaroid disks were set to provide low luminance, a "ready" signal was given, and then the word was flashed onto the screen for 0.2 sec. Luminance-level was increased in approximately 0.005 log-unit steps on succeeding trials by changing angular displacement until correct recognition was achieved. The threshold for each word was measured in one such ascending series for each *S*.

Results. The raw data are in terms of angular displacements of the polaroid disks just large enough for correct recognition. The average of the threshold-determinations for each category was converted into the logarithm of threshold-luminance.

Table I presents the obtained mean log luminance-levels for high- and low-anxious *S*s in each frequency-category, as well as the geometric mean

luminances for reference to the foot-lambert scale. The mean log thresholds are presented graphically in Fig. 1. Inspection of Fig. 1 shows the curves for high- and low-anxious *S*s to be similar in shape, both being negatively accelerated over the range from 0 to 25 reinforcements.

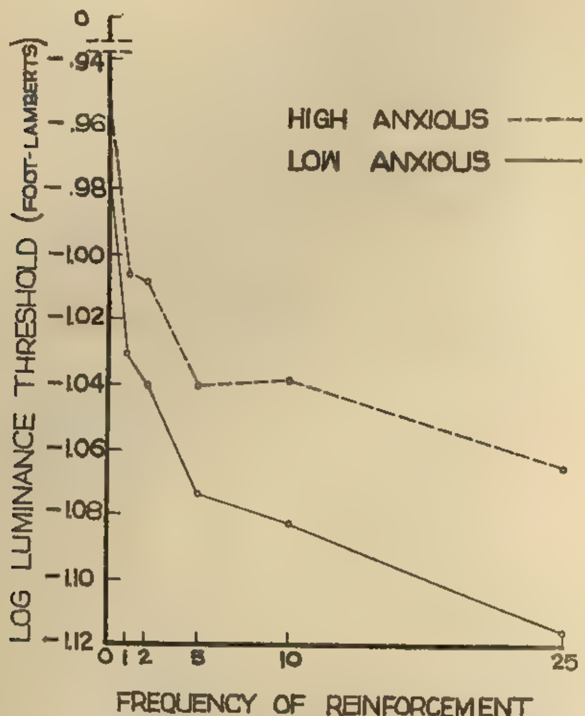

Fig. 1. The Effect of Frequency of Reinforcement and Drive-Level on the Thresholds of Visual Recognition for Nonsense-Words

Further, the mean log-luminance for the group of low-anxious is consistently lower than for the high-anxious, with the differences increasing from the smallest to the largest reinforcement frequency-category.

Analyses of variance of the effects of the level of anxiety and frequency of reinforcement on visual recognitive thresholds of the verbal stimuli were performed. The analyses yielded *F*s significant beyond the 1% level of confidence within each drive-level. Within each drive-group, correlated *t*-tests were computed between adjacent frequencies of reinforcement, and, for both groups, significant differences were found between frequencies of 0 and 1, 2 and 5, 10 and 25, as shown in Table II.

The analysis of variance also focused on the difference between drive-groups to test the second hypothesis. The *F*-value for this comparison approached significance between the 0.05 and 0.10 levels of confidence.

Since unforeseen differences between groups were obtained for words with zero frequency, an effort was made to determine whether there was any initial difference between the groups in threshold-performance. Thresholds from practice trials were averaged. No significant difference ($t = 0.4$, $df = 8$) was found between the two drive-groups.

Discussion. Hypothesis 1 was supported. The reliability of the obtained function is indicated by the fact that substantially the same relationship has been found in two previous studies.[6] The plot describing recognitive responses as a function of the number of reinforcements is similar to the curve that would be expected for instrumental responses as a function of the number of reinforcements. The implications of this demonstration are that general theory of associative learning appears applicable to the analysis of recognitive responses to impoverished stimuli. Under the conditions described for Hypothesis 1, such a formulation was successful. A more adequate test of this notion, however, would necessarily involve a wider sampling of drive-levels.

TABLE II

SIGNIFICANCE-TESTS BETWEEN VISUAL RECOGNITIVE THRESHOLDS FOR ADJACENT FREQUENCIES OF REINFORCEMENT IN TWO ANXIETY-DRIVE GROUPS

Drive-level		Adjacent frequencies 0–1	1–2	2–5	5–10	10–25
low	*t*	4.8	1.0	3.9	.89	3.8
	p	.01	NS	.01	NS	.01
high	*t*	3.6	1.7	2.9	.06	1.8
	p	.01	.10	.01	NS	.05

The failure to support the second and the major hypothesis may be attributed to the inadequate intensity of the drive-conditions. More direct experimental manipulation of this variable using shock or anxiety arousing and allaying instructions of *HA* and *LA Ss* should be attempted. The importance of competition among responses is, however, given weight by the results relative to Hypothesis 2. If competition among correct and incorrect responses under high drive were unimportant, recognition threshold performance would improve and *HA* thresholds would

[6] K. E. Baker and Herman Feldman, Threshold-luminance for recognition in relation to frequency of prior exposure, this JOURNAL, 69, 1956, 278-280; Solomon and Postman, *op. cit.*, 195-201.

be lower than *LA* thresholds, an outcome contrary to the results actually obtained.

Summary. Hypotheses derived from Hull-Spence multiplicative drive-theory were tested. The major hypothesis concerning the relationship between drive-level and recognitive responses to impoverished stimuli was not confirmed, but the difference was in the predicted direction. Further research concerning the competition among responses at high motivational levels seems to be in order.

A RELATION BETWEEN AROUSAL AND PERFORMANCE

By JOHN J. SHERWOOD, Carleton College

According to activational theory, level of performance is related to level of activation or arousal.[1] In general, it is agreed that this relation takes the form of an inverted 'U,' *i.e.* the efficiency of performance is optimal at an intermediate level of activation. Stennett has provided evidence supporting the curvilinear relation.[2] Fiske and Maddi have also reported general support for the curvilinear relation. They summarized research relating degree of arousal to level of performance on various tasks.[3]

Several indices of level of activation or arousal, independent of the performance being measured, have been used, *e.g. EEG,* galvanic skin-response, muscle tonus, pulse- and respiration-rates. The inter-individual correlations between these indices have typically been low.[4] Following Hebb's discussion of the arousal-function of sensory events and his explanation of figure reversals,[5] another index of arousal was available—the number of figure-reversals per unit of time.

This is a report of two experiments which investigate the relation between arousal, as measured by reported rate of figure-reversals, and performance on paired-associate learning tasks.

EXPERIMENT I

Method: (1) Subjects. Fifty-four women undergraduate volunteers served as *Ss*. They were uninformed of the purpose of the research and naïve with respect to paired-associate learning.

(2) Procedure. The experiment was designed as follows: (a) measurement of

---

* Received for publication June 2, 1964.

[1] Elizabeth Duffy, The psychological significance of the concept of "arousal" or "activation," *Psychol. Rev.*, 64, 1957, 265-275; R. B. Malmo, Activation: A neuropsychological dimension, *Psychol. Rev.*, 66, 1959, 367-386.

[2] R. G. Stennett, The relationship of performance level to level of arousal, *J. exp. Psychol.*, 54, 1957, 54-61.

[3] D. W. Fiske and S. R. Maddi, *Functions of Varied Experience*, 1961, 17-56; see also, Malmo, *op. cit.*, 367-386.

[4] Fiske and Maddi, *op. cit.*, 21; see also, J. K. Lacey, Psychological approaches to the evaluation of psychotherapeutic progress and outcome, in E. A. Rubenstein and M. B. Parloff (eds.), *Research in Psychotherapy*, 1959, 160-208; Duffy, *op. cit.*, 269-272.

[5] D. O. Hebb, Drives and the C.N.S. (conceptual nervous system), *Psychol. Rev.*, 62, 1955, 243-254; *A Textbook of Psychology*, 1958, 183-186.

rate of figure-reversals; (b) paired-associate acquisition-trial; (c) retention-trial; and (d) second measure of rate of figure-reversals.

All the *Ss* were familiar with reversible perspectives. *S* was instructed to concentrate on the Necker cube (which was presented on an 8×5-in. card), to attempt to make the cube change perspective, and to say "change" each time it did so. Immediately after the first measure of arousal was taken, brief general instructions on paired-associate learning were given to *S*. The acquisition-trial followed.

Level of arousal was measured by the mean number of figure-reversals of the Necker cube reported by *S* in two 1-min. trials. These two measures correlated, $r = 0.78$ ($p < 0.01$). The range of these mean rates of figure-reversals was 6.5 to 40. The median was 17.

The paired-associate list consisted of 15 two-syllable adjectival pairs with the lowest average similarity values in the Melton and Safier table.[6] The adjectives were typed in large type on 8×5-in. cards, and were presented once at a 2:2 rate. A stimulus-word was followed by a response-word, and both were pronounced aloud. After a 50-sec. interval of rest, each stimulus-word was exposed for 30 sec. while *S* pronounced its associated response-word. The order of presentation was varied to counteract the effects of serial-position. The criterion of performance was the number of correct associations reported after a single trial. Immediately following this retention-test, the second measure of rate of figure-reversals was taken.

Results. The mean rates of figure-reversals from the two trials were used to divide the sample into three groups of 18 *Ss* (high arousal *Ss'* median reversals = 28; medium arousal, 17; and low arousal, 10). The mean retention-scores for these groups were then compared. Mean short term retention-scores as a function of increasing levels of arousal were, 3.0, 4.6, and 5.5. An analysis of variance was performed on these data ($F = 5.59$, $df = 2/51$, $p < 0.01$). The only significant *t*-test comparing the differences between mean retention-scores under the three conditions of arousal was between performance under high *vs.* low arousal ($t = 2.88$, $p < 0.01$).[7]

In addition, rates of figure-reversals and retention-scores were correlated. The Pearson product moment correlation was 0.39 ($p < 0.01$), and rho = 0.44. A scatter diagram showed no curvilinearity.

These results indicated that performance in this study, as measured by short-term retention, was the highest when rate of figure-reversals was the highest. Retention was lowest when rate of figure-reversals was lowest. There was no evidence that performance is most efficient at intermediate levels of arousal. There was also no evidence to contradict this

---

[6] E. R. Hilgard, Methods and procedures in the study of learning, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 548-551.

[7] To control for the possible influence of ability-factors, achievement test-scores for the high, medium, and low arousal groups were compared. There were no significant differences.

proposition, since those levels of arousal at which performance might be expected to decrease might not have been present in this study.

Experiment II

The second experiment was to extend the findings of the first and to test the validity of rate of figure-reversals as a measure of arousal. Due to the low intercorrelations between measures of arousal for different individuals, Fiske and Maddi have suggested designing research to compare different states of the same individual with replications over individuals. They have also suggested that different levels of arousal might be studied by means of subjective reports from the *S*s.[8]

In this experiment, the relation between performance and three levels of arousal for the same *S* was studied. The measures of arousal and performance were the same as those in the previous experiment.

Method: (1) Subjects. Fifty-four women undergraduate volunteers served as *S*s. They were uninformed of the purpose of the research and unconnected with the previous study. From a sample of 75 *S*s, those *S*s who reported the most highly differentiated diurnal patterns of arousal were selected. The larger sample had responded to a questionary by graphically depicting their typical daily changes in levels of arousal, in terms of feelings of "alertness," "being interested in things," "positive emotion," and "being energized."[9]

(2) Procedure. Three appointments were arranged with each *S* on three successive days corresponding to periods of high, medium, and low levels of arousal, as reported on the questionaries. Before the experimental session had begun, each *S* had been asked questions about her subjective feelings of arousal at that time. If her reported level of arousal did not correspond to the scheduled experimental condition, an appointment was made for the following day.

Each *S* was tested three times in a treatments × *S*s design in which both the order of the three conditions of arousal and the effects of the three different tasks were counterbalanced (Lindquist, Type IV).[10] Nine different arousal-task combinations were used and were assigned to the *S*s in a Graeco-Latin square design. Six *S*s were in each sequence.

The measurement-procedure was the same as that of the previous experiment, except that three lists of 12 pairs of nonsense-syllables were used in place of the 15 adjectival pairs. The procedures for acquisition and retention differed from the first experiment in that the syllables were spelled aloud instead of pronounced. Level of arousal was measured by reported rate of figure-reversals of the Necker cube; this was followed by a paired-associate acquisition-trial and a retention-trial; and finally, the second measure of arousal.

The three lists of nonsense-syllables were constructed according to the eight cri-

---

[8] Fiske and Maddi, *op. cit.*, 52-54.

[9] Hebb, *op. cit.*, 1955, 250; Fiske and Maddi, *op. cit.*, 52-53.

[10] E. F. Lindquist, *Design and Analysis of Experiments in Psychology and Education*, 1953, xix, 393.

teria suggested by Melton,[11] and were equated in terms of associational values. All items had associational values of 47% or less.[12]

To control for the effects of practice, *Ss* were given preliminary lists of nonsense-syllables for practice one week prior to the first experimental session. *Ss* were given successive lists of 12 pairs of nonsense-syllables until the number of trials required for criterion (75%) was less than 20. To control for familiarity and the effects of practice with figure-reversals, *Ss* were also given three 1-min. trials with the Necker cube during the preliminary session.

Results. An analysis of variance of the data showed that the only variable which significantly influenced short-term retention was level of arousal. Neither variation in the three learning tasks (criterion-effect) nor order of the three conditions of arousal influenced performance.

A t-test comparing differences between mean retention-scores under the three conditions of arousal showed that there were significant differences between mean retention (3.0) under low arousal and the other two conditions of arousal (mean retention was 5.3 under medium arousal, $t = 2.39$, $p < 0.02$; mean retention was 5.9 under high arousal, $t = 3.02$, $p < 0.01$).[13] There was not, however, a significant difference between retention under conditions of high vs. medium arousal.

One measure of the validity of rate of figure-reversals as a measure of arousal is its relation to *S*'s subjective report of feelings of arousal. A tri-serial correlation between *Ss*' reports of subjective levels of arousal (low, medium, high) and rates of figure-reversals (preacquisition trial) was significant, $r = 0.87$ ($p < 0.01$).[14] This provided some support for rate of figure-reversals as a measure of level of arousal. Before rate of figure-reversals is, however, fully accepted as a useful measure of arousal, it must, of course, be related to more conventional measures—until then it remains only a suggestion.

Summary

Two studies of an exploratory nature investigated the relation between level of arousal, as measured by reported rate of figure-reversals of the Necker cube, and performance, as measured by short-term retention in paired-associate learning. In each study, 54 different undergraduate women served as *Ss*.

In the first experiment, performance was highest under conditions of

[11] Hilgard, *op. cit.*, 540.
[12] Hilgard, *op. cit.*, 541-546.
[13] The median number of figure-reversals under low arousal was 11.5; medium arousal, 20; high arousal, 35.
[14] Nathan Jaspen, Serial correlation, *Psychometrika*, 11, 1946, 23-30.

high arousal and lowest under conditions of low arousal. The second experiment compared performance under different conditions of arousal for the same *S*. Performance was lower under conditions of low arousal than under conditions of medium or high arousal. The validity of rate of figure-reversals as a measure of arousal received support from the finding that there was a significant correlation with the *Ss'* subjective reports of level of arousal.

There was no evidence to support the proposition from activational theory that performance is most efficient at intermediate levels of arousal; neither was there evidence to contradict this proposition, since those levels of arousal where performance might decrease may not have been present in this study.

SOME CONDITIONS OF THE EFFECT OF RELATIVE SIZE ON PERCEIVED RELATIVE DISTANCE

By WILLIAM EPSTEIN and SAMUEL FRANKLIN, University of Kansas

In the absence of other information, a difference between the visual angles subtended by two or more objects may be sufficient to produce a perception of depth. The objects appear to be separated in depth, their perceived relative distances correlated inversely with their relative visual angles, *i.e.* the object which subtends the largest visual angle appears nearest. These observations define the well-known distance-cue called 'relative size.' Evaluations of the relative-size cue are rare. Nevertheless, there seems to be agreement on two points: (1) The effective angular difference is a ratio and not an algebraic difference; and (2) the angular difference leads to perceived depth only when the objects are identically or at least similarly shaped. Since the literature on space-perception contains little experimental evidence on these points,[1] the present study was designed to provide some. Judgments of relative distance were obtained for figures of the same shape and for figures of different shapes. In Experiment I, size-ratio was varied, while the absolute difference remained constant. In Experiment II, absolute size-difference was varied, while the size-ratio remained unaltered.

Experiment I. O sat at a table, his head held motionless by a restraining device. His view was partially obscured by an occluder which restricted him to monocular vision. At a distance of 300 cm. from *O* were two 8 × 8-in. G.E. Electro-luminescent panels. The centers of the panels were separated by a distance of 22.8 cm., and were at approximate eye-level. The intensity of the panels was adjusted to so low a level that the surroundings were not illuminated. The room was completely light-free and all visible surfaces were covered by flat black paint. Under these conditions, only the panels were visible to *O*.

The panels were placed in frames designed to permit *E* to insert an opaque cardboard sheet into the frame, flush against the face of the panel. The standards were prepared by cutting out the desired forms from the cardboard sheet. The luminescent panel shone through the cut-out area.

An additional device was constructed for the purpose of obtaining judgments of distance.[2] This was a 20-in. ruler, calibrated in units of 0.25 in., upon which

* Received for publication June 5, 1964. The research was supported by Grant MH-4153 from the U.S. Public Health Service.

[1] A review of much of the early research on the relative-size cue may be found in M. D. Vernon, *A Further Study of Visual Perception*, 1952, 101-102.

[2] This device was used in an earlier study (William Epstein, Attitudes of judgments and the size-distance invariance hypothesis, *J. exp. Psychol.*, 68, 1963, 78-83).

two markers were mounted. The markers could be moved over the full length of the ruler and the position of the marker could be read directly off the ruler. The markers were raised from the surface of the ruler to prevent O's discriminating the calibration marks tactually.

Standards. The standards were a set of cut-out squares: 6, 4, 3, 2.5, 2, 1, and 0.5 in. on a side, and a set of circles with diameters of 6.77, 4.51, 3.39, 2.82, 2 26, 1.13, and 0.57 in. The standings were paired as follows: square-square, 4 vs. 2, 3 vs. 1, 2.5 vs. 0.5; circle-circle, 4.51 vs. 2.26, 3.38 vs. 1.13, 2.82 vs. 0.57; square-circle, 4 vs. 2.26, 3 vs. 1.13, 2.50 vs. 0.57; square-circle, 2 vs. 4.51, 1 vs. 3.39, 0.5 *vs.* 2.82.

For half of the *O*s, assigned to the square-circle pairings, the square was the larger of the two standards. For the remaining *O*s, the circle was larger than the square.

For each of the four shape-pairings, the ratios of the solid horizontal visual angles varied from 2:1–5:1. The absolute difference was invariant for each shape-pairing: squares = 2 in.; circles = 2.25 in.; square-circles = 2.13 in.

Observers. Sixty students were assigned randomly to six conditions to form six groups of 10 *O*s each.

Procedure. There were six experimental conditions: three size-ratios (2:1, 3:1, 5:1) by two shape-conditions (shapes identical or shapes dissimilar). The same procedure was followed for all conditions. Prior to entering the experimental room, *O* was shown the distance-comparison device. He was instructed concerning its operation, and *E* demonstrated its use. The position of one of the markers was fixed to represent one of the standards. The second marker was free to move and *E* demonstrated how the position of this marker could be adjusted to reproduce various distance-relationships between two standards and an hypothetical *O*. Other aspects of the experimental procedure were also described.

After this preparation *O* was led into the dark room and placed in the desired position. A pair of standards was presented, and *O* made four distance-settings. *O* was instructed "so to adjust the movable marker that its position on the ruler is the same ratio to the stationary marker as the left [right] standard is to the right [left] standard." Two factors were counterbalanced: the lateral position of the fixed marker and the size of the standard (larger or smaller member of the pair) represented by the fixed marker. The judgments were made *tactually.* At no time during the judgments was anything except the standards visible to *O*. This procedure was adopted to avoid the confounding effects of visual-angle relationships between the standards and other objects in the visual field, *e.g.* visual comparison objects.

After his tactual judgments, *O* provided a verbal estimate of the relative distance of the standards, *e.g.* "the right-hand one is three times as distant as the left-hand one."

Results. Table I shows the tactual setting for each condition expressed as a ratio obtained by dividing the setting of the larger standard into the setting of the smaller standard. A 3×2 analysis of variance yielded a significant main effect of ratio ($F = 5.22$, $df. = 2/54$, $P < 0.01$). The effect of shape was not significant ($F < 1$), nor was the shape $\times$ ratio

interaction ($F = 1.40$). Duncan's new range-test showed that the differences between the 5:1 and 2:1 ratios and between the 3:1 and 2:1 ratios were significant ($P < 0.05$). The difference between the 3:1 and 5:1 ratios was not significant. Table I shows that the means for the condition of shape-identity are responsible for the latter finding. An inspection of the individual data reveals two *O*s in the 3:1-identical-shape condition who gave very high ratios (7.63 and 5.35). The mean ratio for this condition, excluding these *O*s, was 2.84 ($N = 8$).

The verbal ratio-estimates are also given in Table I. Analysis of variance showed that only the effect of ratio was significant ($F = 4.36$, $df = 2/54$, $P < 0.05$). Neither the main effect of shape ($F < 1$) nor the shape × ratio interaction ($F < 1$) was significant. Duncan's range-test at the 0.05-level was used to compare the individual mean verbal-estimates. The 5:1 ratio differed significantly from the 2:1 ratio. The remaining comparisons were not significant.

TABLE I

TACTUAL AND VERBAL JUDGMENTS OF RELATIVE DISTANCE EXPRESSED AS RATIOS (EXPERIMENT I)

Angular ratio	Shapes identical				Shapes dissimilar			
	tactual		verbal		tactual		verbal	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
2:1	2.07	0.28	1.95	0.15	2.43	0.81	2.30	0.78
3:1	3.65	1.78	2.90	1.11	2.63	0.85	2.63	0.92
5:1	3.65	1.76	3.55	2.21	3.50	1.29	3.15	1.26

The verbal and tactual judgments were compared by using Wilcoxon's signed rank test. Only the difference for the 3:1 ratio was significant ($P < 0.05$).

Experiment II. The second experiment was the same as the first except in one respect: the ratio was constant (2:1) while the absolute difference varied. The following pairings were used: square-square, 2 vs. 1, 4 vs. 2, 6 vs. 3; circle-circle, 2.25 vs. 1.13, 4.51 vs. 2.26, 6.77 vs. 3.39; square-circle, 1 vs. 2.26, 2 vs. 4.51, 3 vs. 6.77; square-circle, 2 vs. 1.13, 4 vs. 2.26, 6 vs. 3.39.

Results. Table II shows the ratios obtained for the tactual and verbal judgments of relative distance. These data suggest that neither absolute size difference nor shape was a source of variance. The mean ratios for the tactual judgments ranged from 2.12–2.55 with a grand mean of 2.38. For the verbal judgments, the ratios ranged from 2.03–2.37 with a grand mean of 2.17. Analyses of variance yielded no significance main effects or interactions.

Wilcoxon's signed rank-test was used to compare the tactual and verbal

judgments. The differences between the two judgments were significant in the case of the two larger absolute-size differences ($P < 0.02$). Inspection of Tables 1 and II shows that, in each of the three cases where significant differences were obtained, the verbal judgments were closer to the requirements of the visual-angle ratio than were the tactual judgments. It is difficult to interpret this observation, since the order of the tactual and verbal judgments was not counterbalanced.

Discussion. In Experiment I, variations of ratio produced variations in perceived depth despite the invariance of the absolute angular difference. In Experiment II, variations of the absolute difference proved to be ineffective when the ratio was invariant. The conclusion suggested by these results is that the ratio of sizes and not the absolute size-difference is crucial. The pattern of the means in Experiment I suggests that work is needed to determine the psychophysical function relating variations in perceived depth to variations in angular ratio.

TABLE II

TACTUAL AND VERBAL JUDGMENTS OF RELATIVE DISTANCE EXPRESSED AS RATIOS (EXPERIMENT II)

Absolute size-differences (in.) between horizontal angles			Identical Shapes				Dissimilar shapes			
			tactual		verbal		tactual		verbal	
squares	circles	square-circle	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
1	1.12	1.06	2.48	0.91	2.37	0.90	2.38	0.71	2.22	0.20
2	2.26	2.12	2.55	0.81	2.23	0.61	2.34	0.67	2.03	0.38
3	3.38	3.19	2.42	0.60	2.12	0.51	2.12	0.77	2.05	0.93

Neither experiment produced evidence for the importance of shape-similarity. These findings contradict the general opinion that similarity of shape is a necessary condition for the effect of relative size,[3] and they create difficulties for certain explanations of the relative size-effect. For example, in one account it is proposed that *O* initially infers identity of physical size from *observed identity of shape*, and then interprets the difference between the visual angles as a difference in distance.[4] Since

[3] J. J. Gibson, Perception of distance and space in the open air, *Motion Picture Testing and Research*, AAF program, Report No. 7, 1946, 190; W. H. Ittelson, *Visual Space Perception*, 1960, 70-71; J. E. Hochberg and Edward McAlister, Relative size vs. familiar size in the perception of represented depth, this JOURNAL, 68, 1955, 294-296; W. C. Gogel, The size cue to visually perceived distance, *Psychol. Bull.*, in press.

[4] Ittleson, *op. cit.*, 70-71; William Epstein and S. S. Baratz, Relative size in isolation as a stimulus for relative perceived distance, *J. exp. Psychol.*, 67, 1964, 507-513.

there is little reason to believe that *O* would make the initial inference for standards which are not similarly shaped, no depth-effect should be obtained under this condition. Our results contradict any interpretation which makes the effect of relative size a consequence of shape-similarity.

Summary. Two experiments examined the conditions which govern the effect of relative size on relative distance. Two variables were studied: (1) size-ratio vs. absolute size-difference; and (2) shape-similarity. It was found that variations in size-ratio are necessary to produce variations in perceived relative distance. Similarity of shape is unnecessary.

THE EFFECT OF OBSERVATIONAL TECHNIQUE ON BRIGHTNESS-ENHANCEMENT

By Carroll M. Colgan, University of South Carolina

The perceived brightness of an intermittent light is greater at frequencies below fusion than at higher frequencies.[1] Above the critical flicker frequency (*CFF*), the apparent brightness of an intermittent light is not as great as when the same light is applied continuously; instead, the brightness is determined by the ratio of the light-dark phases of the repetitive cycle.[2] Below the *CFF*, the apparent brightness of a flickering light has been found to exceed not only the level expected on the basis of the light-dark ratio (*LDR*) but even the level of the same light applied continuously. This phenomenon has been called "brightness-enhancement" and, by some, the "Brücke effect" after the discoverer.

The present study was designed to investigate the influence of certain instructions on the responses made by the *S*s. Preliminary investigation brought out the fact that judgment of the brightness of a flickering light could be made on the basis of the *on* period alone or on the basis of the *on*- and *off*-periods together. It was decided, therefore, to instruct one group of *S*s to make their judgments on the basis of the *on* period, while the instructions to a second group were in terms of the *on*- and *off*-periods. A third group was given no instructions on this point.

Subjects. The *S*s were 32 men and 4 women. All were either students or faculty members at the University of Florida. Most of the students were advanced graduate students in psychology.

Apparatus. Glow-modulator tubes (Sylvania R-1130B) were used as stimulus-sources. These light-modulating tubes are of the cold-cathode variety. They have the property that the light-output varies in an essentially linear manner with current. The R-1130B tubes give off an almost continuous spectrum at 25 ma. This

* Received for publication July 13, 1964. This paper is adapted from a Master's thesis submitted to the University of Florida. The author is indebted to Dr. Stan E. Wimberly, who directed the research.

[1] S. H. Bartley, Subjective brightness in relation to flash rate and the light-dark ratio, *J. exp. Psychol.*, 23, 1938, 313-319; W. F. Battig, J. F. Voss, and W. J. Brogden, The effect of frequency of intermittence upon perceived brightness, *J. comp. physiol. Psychol.*, 50, 1957, 61-64; S. H. Bartley, Filbert Paczewitz, and Edward Valsi, Brightness enhancement and the stimulus cycle, *J. Psychol.*, 43, 1957, 187-192.

[2] These statements are based on the Talbot-Plateau law (E. G. Boring, *Sensation and Perception in the History of Experimental Psychology*, 1942, 145).

light appears white, but at lower currents the wavelengths in the red predominate. For this reason, red filters were interposed between the two lights and S's eyes. No change in wavelength or perceived hue accompanies a change in current with such filters, and the relation between light-output and current is strictly linear. The manufacturer of these filters is unknown. They were originally selected from a large number of war-surplus filters by means of paired comparisons when held up to a daylight-fluorescent fixture. The two chosen appeared identical.

Spectrophotometric absorption-curves were obtained with a Cary Recording Spectrophotometer (Model 14M).[3] The two filters were found to be essentially equivalent and to be opaque (density greater than the limits of the machine) to wavelengths up to 610 mμ. The curve of transmittance then rose abruptly to a value of 79.4% at 625 mμ. The transmittance of the filters increased to 85% at 640 mμ, with a gradual further increase to 89% at 750 mμ.

The two lights were so mounted in a rectangular optical bench as to illuminate the two red filters mounted in adjacent apertures of 30 mm diameter. A ground-glass screen was placed immediately behind the filters. The light-sources were placed 10 cm. back from this screen with a flat, opaque divider serving to keep the light-beams from overlapping on the screen. The filters were on a level with, and 3 ft. from, the S's eyes. The optical bench was light-tight, and the inside walls were coated with a flat black paint. On the front panel was mounted a viewing hood against which the S placed his head in order to view the illuminated filters inside the bench. Two knobs, which enabled S to vary the brightness of the two lights, were also located on the front panel.

The glow-modulator tubes were part of a circuit which permitted variation of the luminance of the light-source. The circuit was stable and noiseless in operation. It consisted essentially of a positive-bias multivibrator with symmetrical time-constants.[4] The latency of the light-sources was confined to the ionization- and de-ionization-time of the R-1130B glow-modulator tubes.[5]

Procedure. Three groups of 12 Ss each were formed on the basis of the instructions given. Group I received no special instructions. Group II was told to make judgments on the basis of both the *on* and *off* periods of the flickering light. Group III was told to make judgments on the basis of the *on* period alone. Careful questioning indicated that all Ss understood their task.

Other than the special instructions, the procedure was the same for all Ss. After adapting for 10 min. in a partially darkened room, S familiarized himself with the brightness-controls. With both lights on continuously, S made several brightness-matches. His *CFF* was then determined on the basis of the average of one ascending and one descending trial (one light only).

For each brightness-judgment, two lights were presented to S. The left light

[3] This instrument is manufactured by the Applied Physics Corporation of Pasadena, California. The author is indebted to Professor Peyton C. Teague, of the Chemistry Department at the University of South Carolina, who made the measurements.

[4] This electronic switch was used previously by C. M. Colgan (Critical flicker frequency, age, and intelligence, this JOURNAL, 67, 1954, 711-713).

[5] No data on this time were furnished by the manufacturer, but it was stated that the tubes would follow audio-modulation up to 15 kc., which would indicate the time required for ionization or de-ionization to be not more than 1/30,000 sec. Oscillographic inspection of the output substantitates this conclusion within the frequency-range used in this experiment.

provided steady illumination, while the right light was flashing at a rate which was constant for any one trial. *S* was instructed to adjust the brightness of the steady light until it seemed equal to the brightness of the flickering light. The steady light was set at its brightest for the start of each trial. For every trail, the intensity of the flashing light was held constant, with an *LDR* of 1:1. Both lights were off between trials.

There were 20 such trials. Every other *S* received this order of presentation (in flashes-per-second): 32, 26, 20, 6, 14, 6, 7, 13, 3, 18, 10, 18, 16, 11, 9, 4, 12, 5, 12, 8. The order for the alternate *Ss* was: 32, 26, 20, 8, 5, 12, 4, 12, 9, 11, 16, 10, 18, 3, 18, 13, 7, 14, 6, 6. The three highest flash-rates were presented in the first three trials to permit *S* to benefit from what seemed to be the easiest matches demanded of him. Closer inspection of the two orders shows that three reliability-checks were made at 18, 12, and 6 flashes per sec. Furthermore, the check at 18

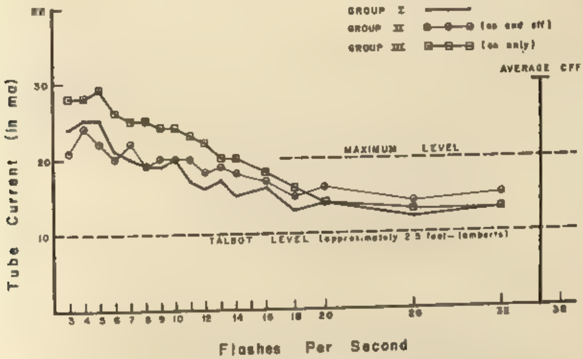

FIG. 1. STEADY INTENSITY REQUIRED FOR EQUAL SUBJECTIVE BRIGHTNESS WITH INTERMITTENT ILLUMINATION

came about halfway through the series for all *Ss*; while the checks at 6 and 12 each occurred early in one series and late in the other.

Results. The results for all three groups are presented in graphic form in Fig. 1, each point representing the mean for 12 *Ss*. Analysis of variance showed that the groups differed significantly ($p < 0.001$), and the difference in judgments on the basis of flash-rate was significant at the same level. As may be seen in Fig. 1, the curves for Groups I and II are rather similar. Orthogonal comparisons showed no significant difference between them, but a significant difference between those two groups compared and Group III ($p < 0.001$) did occur. Reliability checks at 18, 12, and 6 flashes per sec. produced coefficients of 0.47, 0.60, and 0.76, respectively.

Careful and repeated measurements of the light-ouput in ft.-L. were made with a Luckiesh-Taylor Brightness Meter. The line designated "maximum level" in Fig. 1 represents the tube-current of the flashing light

when *on*. If this light had been *on* continuously, rather than intermittently, it would have had a brightness-level of about 5.1 ft.-L., and if the Talbot-Plateau law held as well below fusion as it does above, we would expect the brightness-level of this light, when flashing with an *LDR* of 1:1, to be about 2.5 ft.-L.[6] Ammeter-readings clearly showed 10 ma. for this Talbot level in contrast to 20 ma. for the maximum level.

The vertical axis of the graph in Fig. 1 is presented in terms of the tube-current, because the brightness-measurements are at best only approximate. The difficulty in matching a standard white against a red, as demanded in this case, is well known.[7]

Discussion. Apparent brightness of a flickering light is clearly a function of the flash-rate, even with varied instructions. These results corroborate the findings of Bartley and others that the Talbot-Plateau law does not hold for flash-rates below fusion.

The results of the present study indicate, however, that instructions (or matching techniques) may constitute a significant variable in the study of brightness-enhancement. The *S*s adjust the steady light to a higher intensity for a match based on the *on* period alone, than on both the *on* and *off* periods. The results also indicate that the *S*s who are given no special instructions are likely to make their matches on the basis of both the *on* and *off* periods.

Bartley has used 'brightness-enhancement' to refer to the extent to which the brightness of a flashing light exceeds that of a steady light of the same intensity.[8] In other words, his use of the term does not refer to the extent to which the brightness exceeds the Talbot-level, but to the extent to which it exceeds the brightness expected if the flashing light were on continuously, rather than intermittently. In the present study, this would mean that only where the curves of Fig. 1 exceed 20 ma. would it be proper to speak of enhancement. Since the Talbot-level represents the brightness-level of an intermittent light at all frequencies above fusion, it seems more appropriate to refer to any significant increase in brightness above this level as brightness-enhancement.

Certain differences between this and Bartley's studies should be noted. Bartley used white light interrupted by an episcotister. He refers to a

---

[6] S. H. Bartley, Brightness enhancement in relation to target intensity, *J. Psychol.*, 32, 1951, 57-62.

[7] D. B. Judd, Basic correlates of the visual stimulus, in S. S. Stevens (ed.), *Handbook of Experimental Psychology*, 1951, 817-818.

[8] Bartley, Light adaptation and brightness enhancement, *Percept. mot. Skills*, 7, 1957, 85-92; Edward Valsi, S. H. Bartley, and Charles Bourassa, Further manipulation of brightness enhancement, *J. Psychol.*, 48, 1959, 47-55.

"steady" component of the flickering light.[9] With the electronically modulated square-wave light-output used here, there was no such component present. Bartley has used a variety of targets differing in size, shape, and intensity, but in all cases the intermittent light was adjusted to match the steady light. In the present study, preliminary investigation showed the brightness matches easier to make if the steady light was adjusted, and that procedure was used throughout the experiment.

Summary. Using the method of adjustment, each of 3 groups of 12 *S*s matched a steady light to a flickering light in terms of brightness for each of 17 different flash-rates. The groups differed in instructions. The maximal enhancement occurred at 4-5 c.p.s. for all three groups. There was greater enhancement in a group making its matches on the basis of the *on* period alone than in a group using both *on*- and-*off* periods, or for a group given no special instruction.

[9] Bartley, *Vision*, 1941, 136-138; Valsi, Bartley, and Bourassa, *op. cit.*, 53.

INTRA-PAIR SIMILARITY AS A DETERMINANT OF COMPONENT AND CONFIGURATION DISCRIMINATION

By MICHAEL D. ZEILER and BARBARA J. PAUL, Wellesley College

Teas and Bitterman studied the performance of rats in a discriminative problem involving two pairs of stimuli, each of which was presented in only one of the two possible spatial arrangements (*two-situational problem*).[1] The positive member of one pair was on the left, and the positive member of the second pair was on the right. In such a problem, the animal may come to respond on the basis of the difference between the members of each pair (*component* or *within-pairs* discrimination), or on the basis of the difference between certain global properties of the pairs as wholes (*configurational* or *between-pairs* discrimination). The experiment to be reported provides information about the effect of within-pairs similarity on the amount of within- and between-pairs discrimination which develops in rats confronted with a two-situational problem.

Method: (1) Subjects. The *Ss* were 48 experimentally naïve male Long-Evans rats, 60-70 days old at the start of pretraining. They were divided into six groups designated on the basis of within-pairs stimulus-similarity—minimal (*Min*), medium (*Med*), or maximal (*Max*)—and test-condition—component (*Comp*) or configurational (*Conf*)—as *Min-Comp, Min-Conf, Med-Comp, Med-Conf, Max-Comp,* and *Max-Conf*. The training of Groups *Min-Comp* and *Min-Conf* was completed before that of the other groups was begun.

(2) Apparatus. A two-window modified Lashley jumping stand was used. The square windows, 5.5 in. on a side, were separated by a metal wedge 2 in. wide which extended 2.5 in. in front of the windows. The jumping platform, 7 in. long and 10 in. wide, was enclosed at the sides and back by Plexiglas walls 3.5 in. high and had a Plexiglas roof hinged at the rear. The roof was lifted to place *S* on the stand. The open side of the platform, facing the two windows, had a strip 5.5 in. long and 2.5 in. wide centered in front of each window. Behind the platform was a light-box containing two 100-w. frosted bulbs. A correct response admitted the animals to a box providing access to water, while an incorrect response resulted in a fall into a net 18 in. below the windows. The entire apparatus was painted mid-gray (approximate reflectance = 25%).

Stimulus-doors were made of lightweight sheet-aluminum and were hung on

* Received for publication August 22, 1964. This study was supported in part by Grant MH-08818-01 from the National Institute of Mental Health. The authors are grateful to Aida Price and Shirley Wang for their assistance in taking the data.

[1] D. C. Teas and M. E. Bitterman, Perceptual organization in the rat, *Psychol. Rev.*, 59, 1952, 130-140.

hooks attached at the rear of the windows, 1 in. above the top of the opening. A dowel could be inserted behind either door to prevent it from opening.

Stimuli. The discriminanda were gray doors differing in brightness and striped doors (.75-in. wide black and white stripes) differing in the degree to which the stripes were tilted from the horizontal. The two pairs of doors used for *Min, Med,* and *Max* groups are listed in Table I. The pretraining stimuli were mid-gray doors (approximate reflectance = 25%).

Pretraining. A 23.5 hr. water-deprivation schedule was used. After three days of gentling, *S*s were placed individually in the watering-box and then allowed to walk through open windows from the jumping platform to the watering-box. The animals were trained to jump from gradually increasing distances to fully open windows and then to unobstructed mid-gray doors. The animals were guided in jumping to both doors. Pretraining ended when *S* jumped readily for 10 trials on each of two days from 9 in. to these doors. Brief access to water was permitted on each trial.

Situational training. Two-situational training followed pretraining. The *S*s were differentially rewarded for responding in one direction to a pair of grays and in

TABLE I

STIMULUS-PAIRS

Condition	Percentage reflectance (grays)		Degrees of tilt from horizontal (stripes)	
	Stimulus 1	Stimulus 2	Stimulus 1	Stimulus 2
min	3	76	0	90
med	11	43	45	135
max	18	31	22.5	157.5

the opposite direction to a pair of stripes. Counterbalancing led to eight subproblems for each group to each of which one animal was assigned randomly. Whether grays or stripes were presented on a given trial was predetermined randomly with the stipulation that each problem appear twice in every block of four trials.

Each rat received 12 trials per day. The intertrial interval was the time required for *S* either to have several licks at the spout of the water bottle or to be picked up from the net, for the appropriate doors to be hung, and the dowel positioned for the next trial. The *S* spent the latter phases of the intertrial interval in a carrying cage. On the first day, a guidance technique was used: *S* was given one free jump on each trial, and, if incorrect, was guided to the correct side. On all subsequent days, a noncorrection technique was employed. Criterion was 11 of 12 correct choices each of two successive days.

Situational test. On the day following criterion in the two-situational problem, the four-situational test was begun; that is, the stimuli comprising each pair were presented in both possible spatial arrangements. In Groups *Min-Comp, Med-Comp,* and *Max-Comp,* the same doors that were positive in the two-situational problem continued to be reinforced regardless of their spatial position. For Groups *Min-Conf, Med-Conf,* and *Max-Conf,* the same spatial position that was positive in training for a given stimulus-pair continued to be reinforced regardless of which door occupied that position.

The order of presentation of the stimulus-pairs was random with the stipulation that each of the four possible situations appear in every block of four trials. The number of trials per day (12), the procedure in the intertrial interval, and criterion (11 of 12 correct choices on each of two successive days) were identical with two-situational training. A noncorrection technique was used throughout.

Results. Two-situational problem. The mean errors to criterion did not differ for the six groups ($F < 1.00$). The over-all mean number of errors to criterion was 28.7.

Four-situational problem. The number of *S*s in each group that re-

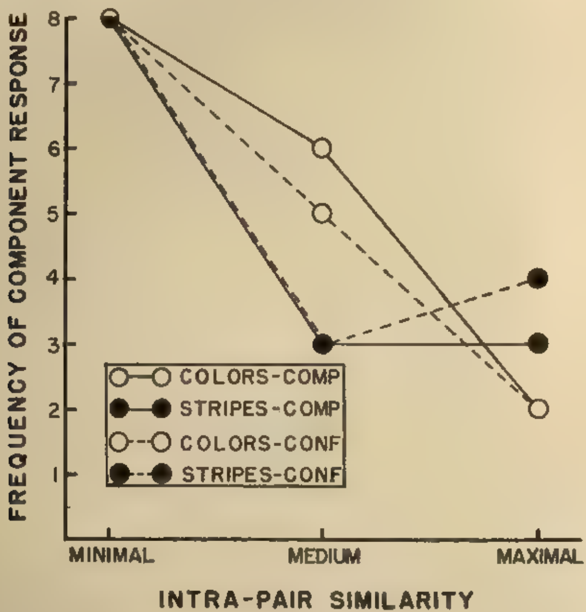

FIG. 1. FREQUENCY OF COMPONENT-RESPONSE FOR THE FIRST TRIAL OF FOUR-SITUATIONAL TEST ON WHICH THE STIMULI APPEARED IN THE NEW SPATIAL ORIENTATIONS Data are presented separately for the gray and stripe problems and for the type of test-problem, component (*Comp*) and configurational (*Conf*.).

sponded to the component (door) previously reinforced in the two-situational problem when the pair was encountered in altered spatial arrangement for the first time in the four-situational problem is shown separately for the grays and stripes in Fig. 1. All of the *S*s in Groups *Min-Comp* and *Min-Conf* chose the previously reinforced component in both pairs. The only other non-random outcome in the data for the initial presentations

indicated configurational transfer with the gray pairs for Groups *Max-Comp* and *Max-Conf* combined ($\chi^2 = 4.00$, $df = 1$, $p < 0.05$).

While seven of the eight *S*s in Group *Min-Comp* learned their test-problem without a single error, and the other *S* made only two errors, no *S* in Group *Min-Conf* reached criterion even on a single day. Five of these eight *S*s never surpassed the level of 50% correct; three *S*s persisted in response to the previously reinforced components, while the other two fell into rigid position-preferences. The remaining *S*s responded at the level of 75% correct in their last day on the test-problem; two learned the appropriate response to the stripes, and one learned the response to the grays.

Groups *Med-Comp* and *Med-Conf* did not differ significantly in the mean number of errors to criterion ($t = 0.28$, $df = 14$, $p > 0.5$). The combined mean errors to criterion was 11.25. In each of these groups, all of the *S*s reached criterion within 11 days and five *S*s made 10 or fewer errors. Errors for the other three *S*s were 14, 16, and 26 in Group *Med-Comp* and 19, 25, and 33 in Group *Med-Conf*. Most of the errors for these six *S*s were on one problem which the other mastered with six or fewer mistakes. Two of the three *S*s in the component group missed more stripes than grays, while two of these three *S*s in the configuration group erred more frequently on the grays.

All of the *S*s in Group *Max-Conf* and one *S* in *Max-Comp* reached criterion by Day 11, while four *S*s in Group *Max-Comp* failed to reach criterion by Day 12 when their training was stopped. The other three *Max-Comp* *S*s learned in 15, 18, and 21 days, respectively, which indicated that the two members of each stimulus pair were discriminably different. The mean error in the first 12 days of training for Group *Max-Comp* was 45.8 and for Group *Max-Conf* only 6.4 ($t = 7.54$, $df = 14$, $p < 0.001$). In Group *Max-Conf*, five *S*s made three or fewer errors and the other *S*s 12-18 errors. Of the latter, two made virtually all of their mistakes on the stripe problem, while the third had about equal difficulty with the grays and stripes. In Group *Max-Comp*, all the *S*s made more than 30 errors most of which were on the gray problem; five *S*s made fewer than seven mistakes with the stripes.

Discussion. The relationship between intra-pair similarity and the basis of solution of the two-situational problem is clear. Component-learning occurred with minimal similarity, configurational learning with maximal similarity, and both forms of learning appeared in the middle-similarity range. Since for many *S*s the majority of errors in the four-situational test-problem were with one stimulus-pair, it seemed that animals simulta-

neously could learn one of the discriminations on the basis of components and the other on the basis of configurational properties.

The results provide evidence that either component or configurational learning may have functional priority in the two-situational problem and, therefore, support a dual-process analysis of discriminative behavior.[2] The component problem is solved on the basis of selective orientation to the specific reinforced stimuli (*i.e.* approach vertical stripes and white). In a configurational discrimination, the effective cue is some unitary property of the stimulus-pairs to which *S* responds directionally (*i.e.* jump right to stripes, jump left to grays). The independence of these processes is emphasized by the current finding that each pair in a two-situational discrimination may be learned on a different basis (*i.e.* approach the stripes that is tilted in a given direction and jump left to a pair of grays).

Summary

The dominance of component and configurational learning in rats was found to be a function of intra-pair similarity. When this similarity was minimal, there was perfect component learning. With intermediate levels of intra-pair similarity, equal learning of both types resulted, while discriminably different but maximally similar stimuli were learned configurationally.

[2] M. E. Bitterman and Jerome Wodinsky, Simultaneous and successive discrimination, *Psychol. Rev.*, 60, 1953, 371-376; M. E. Bitterman, D. W. Tyler, and C. B. Elam, Simultaneous and successive discrimination under identical stimulating conditions, this JOURNAL, 68, 1955, 237-248.

THE INFLUENCE OF CONTEXTUAL VARIATION ON THE DIFFERENTIATION OF PARTS FROM WHOLES

CYNTHIA TURNURE, Yale University, and LISE WALLACH, University of North Carolina

It has been asserted repeatedly, without much in the way of evidence, that variation in the context of a stimulus facilitates differentiation.[1] The present experiments were designed to provide such evidence.

Method. Each of the *S*s of Experiment I (48 Wellesley undergraduates) was given a pack of 15 3 × 5-in. white cards, on each of which was drawn in black ink one of the training stimuli shown in Fig. 1. *E* gave the following instructions:

> I would like you to sort these cards into three equal groups. There are 15 cards and you are to put 5 cards in each group. You may sort them any way you wish—there is no right way to do it—and you may take as long as you need. Remember to end up with five cards in each group.

For half the *S*s (*A*-in-Variation Group), Stimulus-Part *A* occurred in five different figures (in each one with a different concomitant), while Stimulus-Part *B* occurred five times in a single figure (always with the same concomitant). For the remaining *S*s (*B*-in-Variation Group), Part *B* occurred in five different figures, while Part *A* recurred five times in a single context. Each of the six *A*-figures and six *B*-figures reproduced in Fig. 1 served as the stimulus that was duplicated five times for four *S*s. The 10 other figures used were the ones that did not include the concomitant repeated with Part *A* or *B*. Thus, for example, four *S*s in the *B*-in-Variation Group received five cards with A_1 and cards with all the *B*-figures and *C*-figures except B_1 and C_1. The *C*-figures were included to prevent the *S*s from simply putting their identical cards together without considering any similarities and differences among the other figures. The 15 cards composing each pack were arranged in a randomized order, a different order being used for each pack.

After each *S* sorted the 15 training stimuli, she was shown a card with Part *A* and Part *B* drawn as separate figures (the Test Stimuli of Fig. 1) and asked, "Which of these two figures do you think you have seen before?" If *S* indicated that she had seen both of them, she was asked, "Which one seems more familiar to you?" It was predicted that the *S*s in the *A*-in-Variation Group would be more likely to choose Part *A*, while *S*s in the *B*-in-Variation Group would be more likely to choose Part *B*.

The design of Experiment II was the same as that of Experiment I, except in three respects: (1) Exposure-time was controlled; (2) the data were used for only those *S*s who believed that they had seen one, and only one, of the test-figures before; and (3) two different training procedures were used.

* Received for publication September 24, 1964. This research was initiated at Wellesley College and completed at Yale University.

[1] William James, *Psychology: The Briefer Course*, 1892, 249-251; George Humphrey, *Thinking*, 1951, 295.

In Experiment II, the training stimuli were presented serially, each figure being exposed for 1 sec., in a different randomized order for each *S*. Before the series, half the 24 *S*s in each of the two groups were given the following instructions for training in examining:

I wish you to look carefully at a series of abstract figures and to get to know them as well as you can. There are 15 figures, and you will be shown each one for

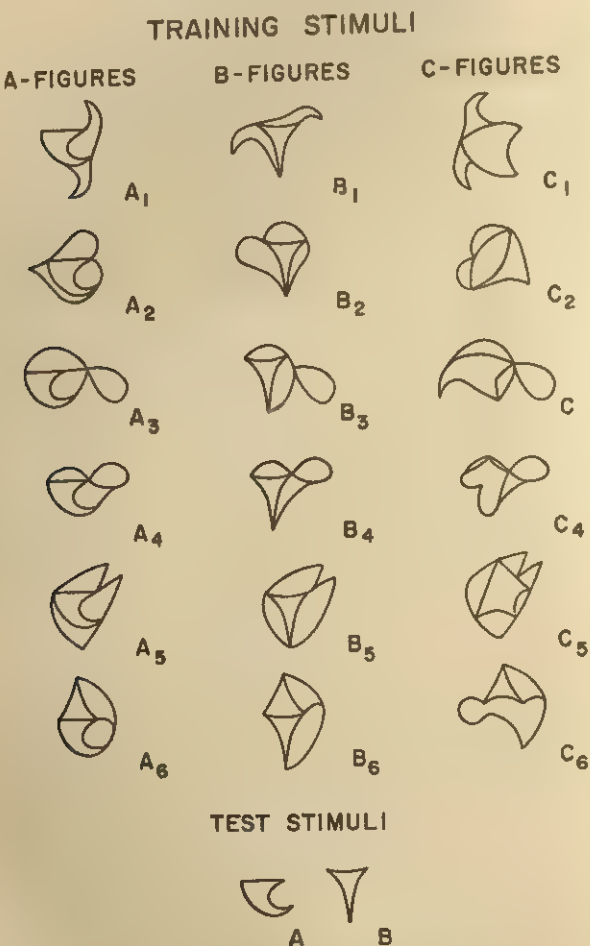

FIG. 1. TRAINING AND TEST-STIMULI

1 sec. Afterward, I will test your memory for the figures, so you should become as familiar with all of them as possible. You may have a chance to get to know some figures better than others, because some may be repeated; however, each figure is important, and you should try to learn as much about each one as you possibly can.

The other half of the *Ss* in each group received the following instructions for training in Concept-Attainment:

I wish you to look carefully at a series of abstract figures, and to get to know them as well as you can. Each figure has a nonsense-name below it. There are 3 names and 15 figures, so each name is the correct name for 5 figures. You will be shown each figure-name combination for 1 sec. Afterward, I will test your memory for the figures and the names that go with them, so you should become as familiar as possible with them all. You may have a chance to get to know some of the figures and their names better than others, because some combinations may be repeated; however, each figure-name combination is important and you should try to learn as much about each one as you possibly can.

One of the three nonsense-syllables, *QIH*, *XUD*, and *ZEJ*, was randomly assigned to be presented below each stimulus-card with a given *S*'s identical figures, another with the five figures containing the part occurring in variation, and the third with the five remaining figures.

After each *S* (a Yale undergraduate) viewed his 15 figures, he was shown the recognitive test-card and asked, "Do you think either one of these forms has appeared in any of the figures you have just looked at?" If *S* answered "Yes" he was further questioned as follows: "Which one? Why did you remember it? Do you think the other form might also have appeared?"

TABLE I

CHOICES OF ALL *Ss* ON THE RECOGNITIVE TEST IN EXPERIMENT I

Group	Part A	Part B	Total
A-in-Variation	17	7	24
B-in-Variation	9	15	24
Total	26	22	48

The data for only those *Ss* who answered "Yes" to the first question above and who also gave one of the following responses when asked, "Do you think the other form might also have appeared?": "No," "Don't know," "Don't recall," "Couldn't say," "Can't tell," "Don't remember," or "Don't remember it," were used. To obtain the 48 *Ss* meeting this criterion, 125 *Ss* were tested. The two investigators independently judged the coded protocols, with agreement on 98% of the cases. The remaining cases were resolved by discussion.

Results. The choices of all *Ss* on the recognitive test in Experiment I are shown in Table I. Comparing the frequencies with which *A*-in-Variation *Ss* and *B*-in-Variation *Ss* chose Part *A* and Part *B*, we find that $\chi^2 = 4.11$, $p < 0.05$ with 1 *df*. If the 10 *Ss* who had some difficulty in choosing which test-stimulus had been seen before are excluded, $\chi^2 = 8.97$, $p < 0.005$ with 1 *df*.

Table II shows the frequencies with which *A*-in-Variation *Ss* and *B*-in-Variation *Ss* with the two kinds of training chose Part *A* vs. Part *B* in Experiment II. Each part was chosen more frequently by the group for which it had occurred in variation, with $\chi^2 = 6.80$, $p < 0.01$ with 1 *df*, for *Ss* in both training groups combined. There also was significantly more frequent

recognition of the part occurring in variation among *Ss* with Examination-Training alone ($\chi^2 = 4.29$, $p < 0.05$ with $1df$) though not among *Ss* with Concept-Attainment Training alone ($\chi^2 = 1.50$, $0.20 > P < 0.30$ with 1 df). The difference between the frequencies with which *Ss* in the two training groups chose the part that had occurred in variation is, however, not significant ($\chi^2 = 0.10$).

Discussion. The results appear to demonstrate the influence of contextual variation on the differentiation of parts from wholes. This outcome may be viewed as an illustration of the tendency for cognition to be as economical as possible, the evidence for which has been reviewed by Hochberg.[2] The cognition of repeated stimuli or invariances would require the differentiation of parts when they occur in variation, but not when they occur always in the same context.

TABLE II

CHOICES ON THE RECOGNITIVE TEST OF 48 *Ss* MEETING THE CRITERION IN EXPERIMENT II

Group	*A*-in-Variation		*B*-in-Variation		Total
	Part A	Part B	Part A	Part B	
Concept-Attainment	8	4	4	8	24
Examination-Training	8	4	2	10	24
Total	24		24		48

This interpretation of course leaves unanswered the question of *how* differentiation is facilitated by variation. One possible factor may be the difference in ease of differentiation of a part in different wholes. The probability that a certain common part will be unusually easy to differentiate in one of a number of different figures is greater than the probability that it will be unusually easy in any given figure. Comparison among the proportions of *Ss* trained with each figure as their repeated figure who then chose the critical part of that figure as familiar suggests, however, that the results of the present experiments are not due simply to such differences in ease of differentiation in the various contexts. The proportion choosing the critical part in any repeated figure is never larger than $5/8$, and the proportions show no tendency to skew.

Two other explanations may be offered: One is based on the alleged arousing effects of varied as opposed to monotonous stimulation.[3] The

[2] J. E. Hochberg, Effects of the Gestalt revolution: The Cornell symposium on perception, *Psychol. Rev.*, 64, 1957, 73-84.

[3] D. W. Fiske and S. R. Maddi, *Functions of Varied Experience*, 1961, 11-56.

other is provided by Hebb's cell-assembly theory,[4] according to which a stimulus is more likely to be recognized if it has previously occurred in a varied context rather than always in the same context. In the former case, a separate assembly of the cortical cells excited by that stimulus would have developed and would be activated by that stimulus presented separately; in the latter case, however, there would be only one assembly whose activation would depend on the presence of both the stimulus and its context.

Summary. The results of two experiments support the hypothesis that a stimulus which is presented in varying contexts is more likely to be differentiated from its surroundings and recognized in subsequent situations than a stimulus which always occurs in the same context.

---

[4] D. O. Hebb, *The Organization of Behavior: A Neuropsychological Theory,* 1949, 69-74, 88-90.

GAMMA-MOVEMENT AND THE PUPILLARY REFLEX

By I. T. Baldwin, University of Kentucky and
Roger W. Tinsley, Centre College

Bartley distinguished six different kinds of gamma-movement and attributed all of them to a spatio-temporal distribution of the retinal image resulting from differences in reaction-time of rods and cones and differences in intensity of stimulation from the center toward the periphery of the retina.[1] He purported to discredit earlier theories which attributed these phenomena to central processes. The earlier explanations were advanced by the Gestalt psychologists, who saw gamma-movement simply as a psychological phenomenon concomitant with any figure emerging on a contrasting ground.

The present experiment is concerned with only one of the six movements described by Bartley: the apparent centrifugal movement of the contour of a lighted disk as the light emerges, and the opposite movement as the light decays. Within specified ranges of stimulus-size and cyclical rate of presentation, the disk will appear to expand and contract. The experiment was designed to test a still more parsimonious explanation of this phenomenon than either of those already considered—an explanation in terms of the pupillary reflex.

It was found that a sharp and accurate representation of the pupil could be obtained by placing a small hand-mirror over the eye in a semi-dark room. The pupil-reflection was seen as a bright disk on an amorphous dark field. Very slight changes of light-intensity near the eye resulted in easily discernible changes in pupil-size. The first significant finding was that turning on and off a small light-source near the eye at the optimal speed for gamma-movement revealed a direct relationship between emergence of the light and pupil-size. This relation was the opposite of what one might expect, since the pupil normally closes as light-intensity increases. It was decided that the observed effect was due to a lag in the reflex which reversed its relationship to the emerging and decaying light-source. An earlier theoretical attempt by the authors to relate contour-pulsations of the stimulus

* Received for publication August 17, 1964. The authors acknowledge with thanks the assistance of Joan Lee and John Donahoe.

[1] S. H. Bartley, *Vision: A Study of Its Basis*, 1941, 158-165; *Principles of Perception*, 1958, 264-266.

to diffraction and spherical aberration had failed because spherical aberration normally increases directly with pupil-size. For this reason, one would expect diffraction and spherical aberration to decrease rather than increase with emergence of a lighted stimulus. With the discovery of the direct relationship between pupil-expansion and apparent outward movement of the stimulus boundary, however, spherical aberration could be employed as a determining variable, since those areas of the lens responsible for spherical aberration are alternately exposed.

The present experiment was designed to observe the effect of eliminating the pupillary reflex on the perception of gamma-movement. The purpose of the apparatus and method was to immobilize the pupil in a constricted position.

Apparatus. The stimulus-object was a box, 18 in. square and 12 in. deep, enclosed on all sides. One side of the box was removed and replaced by a facing of black posterboard which had been pierced by a circular hole 2 in. in diameter. Four sheets of onion-skin paper were pasted over the back of the hole. A 40-w. incandescent bulb mounted on a base was placed in one corner of the box, out of line with the hole. A normally-open lever-type microswitch was placed in circuit with the light and mounted beneath the shaft of a continuously adjustable variable-speed motor. The shaft of the motor was fitted into one end of an elliptical cam which depressed the microswitch during one half of each revolution of the shaft. Frequency of light-flashes varied directly with the speed of the motor.

Another box, open on one side and lined with aluminum foil, was placed on a table 5 ft. from the stimulus-box. A 300-w. incandescent bulb was mounted in the top of the box. A horizontal hole, 8 × 2 in., was cut beneath the bulb on the side of the box facing the stimulus. A dimmer, placed in the circuit with the 300-w bulb, was constructed by filling a 4-in.-deep Pyrex dish with salt water. One end of the broken light-circuit was soldered to a copper penny and placed in the bottom of the dish. The other end also was soldered to a penny and mounted on an insulated shaft which was driven by a set of gears mounted above the dish. By turning a knob attached to the gears, the shaft could be moved smoothly from above the water level down to a point where the pennies made contact. The closer the pennies came to each other the brighter the light became. To obtain a measure of brightness, a voltmeter graduated in steps of 2 v. was connected between the 300-w. bulb and the dimmer.

O sat behind the foil-lined box, leaning forward, with his head inside the box. His eyes were in line with the 8 × 2-in. aperture, facing the stimulus. He was provided with a push-button operated signal light. Preliminary testing indicated semi-darkness was preferable either to full light or total-darkness for the room in which the experiment was conducted.

Instructions. After gamma-movement was explained to *O*, he was asked to adjust the speed of the motor until he observed the most pronounced movement. The optimal exposure-interval was approximately 250 m sec. Then, as *O* fixated the stimulus, the dimmer-switch was turned up by *E*, who told *O* to press the signal-button when absolutely sure he could no longer detect pulsations of the stimulus-boundary. Following this, *O* was given three trials. The voltmeter-readings, taken

when the signal-light was on, were recorded. When three trials were over, *O* was asked to continue to fixate the stimulus while *E* observed his pupil. In all cases, *O*'s pupil was found to have reached full constriction at the point at which he reported disappearance of gamma-movement, since no further constriction resulted from a continued increase of intensity to full strength of the 300-w. bulb.

Observers. Nine *O*s, four men and five women, were used. Their ages ranged from 17-44 yr. with a mean of 29.7 yr. All were above average in intelligence as indicated by superior academic achievement.

Results. All nine *O*s reported perception of the kind of gamma-movement under investigation and its disappearance as the intensity of the 300-w. bulb was increased. Table I shows much less variation over trials than among *O*s.

The results appear to support the hypothesis that fluctuation of figural contours under the conditions described is a function of the pupillary

TABLE I

VOLTAGE AT WHICH MOVEMENT DISAPPEARED

Observer	Trial 1	Trial 2	Trial 3
1	70	60	60
2	55	58	56
3	66	64	64
4	50	50	48
5	50	50	48
6	50	52	50
7	60	56	56
8	62	64	64
Mean	57.2	56.7	55.7

reflex. It seems that the sharpness of contours under expansion or contraction is a crucial consideration. If the effect were due to a sloping gradient of intensities, as Bartley has suggested, one would expect poor definition of boundaries in the expanded condition rather than the sharp boundaries actually observed. If changes in diffraction were responsible, maintenance of definition could be expected, since fluctuations as slight as those observed here would fall well within the eye's depth of focus. Krauskopf's results would support this interpretation.[2] He showed imagery to be approximately constant in quality over pupil-sizes varying from 3-5 mm.; in the present experiment, the constricted pupil position always was less than 5 mm. in diameter.

One possible objection to the present finding is that a bright light shining into the eyes of *O* would obliterate the minimal stimuli postulated by Bartley. To answer such an objection, the present investigators elimi-

[2] John Krauskopf, Light distributions in human retinal images, *J. opt. Soc. Amer.*, 52, 1962, 1046-1050.

nated the 300-w. bulb and observed the stimulus through an artificial pupil (a pinhole in a card). Fluctuations of the contour disappeared under these conditions, too.

It should also be mentioned that similar results are obtained either with a 25-, a 40-, or a 60-w. bulb used to light the stimulus. Since current for the bulbs was supplied directly from a 110-v. A.C. line, no other control was needed. It also has been suggested that glare from the stimulus may be a significant factor. It was decided in pilot trials to conduct the experiment in a semi-darkened room for several reasons. In total darkness, the emerging stimulus, appearing from nowhere, tended not only to glare but to create discomfort, and points of reference for localization of the stimulus were lost in the off-stage. It was felt that in a fully lighted room, peripheral stimuli might be distracting. In a semi-darkened room, the stimulus, arising from a somewhat amorphous gray field, was more comfortable to observe, *O* had no difficulty keeping it oriented, and there was no observable glare as it emerged. It should be added, however, that results regarding the cessation of gamma-movement are the same under all three ambient-light conditions.

Another possible objection is that the apparatus and method were somewhat crude for the investigation of anything so delicate as the human visual system. It is reasonable to think, however, that all significant variables were sufficiently controlled to demonstrate at least the gross aspects of the phenomenon under investigation. Certainly, the findings are statistically reliable, although it is recognized, of course, that refinements in technique would yield more precise measurements.

Summary. The relationship of the pupillary reflex to fluctuating contours in gamma-movement was investigated. Constriction and immobilization of the pupil by a bright light during the observation of gamma-movement resulted in the disappearance of the contour-fluctuations. The results bear on S. H. Bartley's theory that all kinds of gamma movement are due to the spatio-temporal distribution of the retinal image.

EYE-MOVEMENTS AND THE AUTOKINETIC ILLUSION

By Richard S. Lehman, University of Colorado

One of the earliest theories of autokinetic movement was proposed by Hoppe.[1] He suggested that the illusion was caused by unconscious eye-movements. Guilford and Dallenbach reviewed the evidence against Hoppe's theory, photographed *Ss'* eyes during the illusion, and concluded that eye-movements are not essential features of the autokinetic illusion.[2]

Skolnick reviewed the existing criticisms leveled against the eye-movement hypothesis and concluded that none of the previous research fully invalidated Hoppe's theory.[3] Skolnick went on to show (1) that autokinesis could be induced by introducing rotational or caloric nystagmus, and that such induced movement could not be distinguished from normal autokinetic movement; and (2) that *E* could reliably predict the direction of autokinetic movement by direct observation of *S*'s eyes.

Matin and MacKinnon have recently shown the importance of eye-movements in the autokinetic phenomenon.[4] Using stimuli stabilized on the retina against horizontal movements, they found that the horizontal component of autokinetic movement disappeared almost completely. It thus appears that eye-movements are an essential feature of the autokinetic illusion.

This paper presents some further findings that implicate eye-movements as causative factors in the autokinetic illusion. Specifically, the work was designed to answer two questions: First, are eye-movements related to the initiation of autokinetic movement? Secondly, are eye-movements independent, in a probability-sense, of autokinetic movements?

Method. An American Optical Company ophthalmograph was adapted for use with infrared light. Further modifications of the instrument made possible the presentation of an autokinetic point-source of white light during the recording of eye-movements with the infrared light, and allowed *S* to place an indication on the

* Received for publication May 14, 1964. This paper is based upon a Master's thesis submitted to the University of Colorado in 1963. The advice of Michael Wertheimer and Daniel E. Bailey is acknowledged with thanks.

[1] J. I. Hoppe, *Die Scheinbewegung*, 1879, 84.

[2] J. P. Guilford and K. M. Dallenbach, A study of the autokinetic sensation, this Journal, 40, 1926, 83-91.

[3] Alec Skolnick, The role of eye movements in the autokinetic phenomenon, *J. exp. Psychol.*, 26, 1940, 373-393.

[4] L. Matin and G. E. MacKinnon, Autokinetic movement: Selective manipulation of directional components by image stabilization, *Science*, 143, 1964, 147-148.

film of all starts, stops, and changes in direction of the autokinetic movement. Such changes in autokinetic movement will be referred to hereafter as *autokinetic reports.* Each of seven *Ss* fixated the autokinetic point for approximately 90 sec., indicating all starts, stops, and changes in direction of the illusory movement. Head-movements were minimized by the chin- and forehead-rests and the temple-clamps of the ophthalmograph.

Results. For each *S*, a frequency-distribution of eye-movement size was obtained. Saccades, in any direction, of sufficient size to place them above the 70th percentile in the distribution were arbitrarily called *large.* For all seven *Ss*, the first report of autokinetic movement was preceded within ¼ sec. by a large eye-movement; the binomial probability of all seven *Ss* reporting such movement is 0.0078. Thus, the initiation of autokinetic movement is clearly related to eye-movements and the first question is answered affirmatively.

The answer to the second question comes in the form of a test of the null hypothesis that eye-movements are unrelated to autokinetic movements. The null hypothesis states that autokinetic report is independent of eye-movements. If the two are independent, then P (autokinetic report) $= P$ (autokinetic report after large eye-movement). The procedure for testing this hypothesis is based upon the normal approximation to the binomial. The inspection-period was broken down into consecutive 1-sec. intervals, and each interval treated as a binomial trial. For the group as a whole, the first probability was 0.134, the second 0.550, and the difference was significant beyond the 1% level.[5]

The null hypothesis may therefore be rejected; autokinetic movement is not independent of eye-movement.

Discussion. It is recognized that the ophthalmograph is not the most accurate method of recording and measuring eye-movements.[6] The method was, however, adequate for the purposes of this study. If the work were repeated with a more refined recording and measuring procedure, the results undoubtedly would be even more clear-cut.

In spite of the procedural weakness due to the use of the ophthalmograph, the research reported here offers further support for the idea that eye-movements are necessary to the autokinetic illusion; they are not necessarily the sole cause. Although no further suggestion as to mechanism

[5] F. L. Wolf, *Elements of Probability and Statistics*, 1962, 228. The probabilities for each *S* were in the predicted direction. A t-test showed a significant difference between the two sets of probabilities ($t = -13.51$, $df = 12$, $p < 0.001$).

[6] L. A. Riggs, J. C. Armington, and Floyd Ratliff, Motions of the retinal image during fixation, *J. opt. Soc. Amer.*, 44, 1954, 315-321.

can be made, it should be noted that there is considerable evidence for some sort of central involvement. A purely peripheral explanation of the phenomenon cannot, for example, account for the ability of prior figural inspection to reduce the extent of autokinetic movement.[7]

Summary. The research reported demonstrates two relationships between eye-movements and the autokinetic illusion: (1) eye-movements are consistently related to the onset of autokinetic movement; and (2) eye-movements are consistently related to starts, stops, and changes in direction of autokinetic movement.

---

[7] See, for example, R. S. Crutchfield and Ward Edwards, The effect of a fixated figure on autokinetic movement, *J. exp. Psychol.*, 39, 1949, 561-568; Edwards and Crutchfield, Differential reduction of autokinetic movement by a fixated figure, *ibid.*, 42, 1952, 25-31; J. R. Royce, W. R. Stayton, and R. G. Kinkade, Experimental reduction of autokinetic movement, this JOURNAL, 75, 1962, 221-231.

THE EFFECTS OF REPEATED CONFINEMENT ON COGNITIVE PERFORMANCE

By PETER SUEDFELD, JACK VERNON, JOHN T. STUBBS, and MARVIN KARLINS, Princeton University

While subjects generally report cognitive disturbances during sensory deprivation, objective tests have given inconsistent results.[1] A previous study indicated that this contradiction may have been due to the use of clear and structured test-problems, while subjective reports typically dealt with unstructured tasks. Using an oral TAT-technique, and emphasizing the desirability of long and detailed responses, we found that sensorially deprived and socially isolated *S*s (the *SD* Group) told much shorter stories after than before 24 hr. of confinement; isolated but not otherwise deprived *S*s (the *SI* Group) increased their story-length; and there was no difference in the stories of unconfined control *S*s (the *C* Group). These results were interpreted as showing heightened social drive after isolation, which, however, was outweighed by a cognitive decrement (lessened ability to concentrate) produced by sensory deprivation.

The present study, besides replicating the earlier one, was designed to determine whether repeated confinement would reduce these effects. There is evidence that a second period of sensory deprivation results in fewer reports of unusual experiences,[2] and we wished to see whether the cognitive and affective consequences of a single *SD* or *SI* experience would disappear after a second session.

Subjects. Our *S*s were 58 college men, each of whom had volunteered to lie on a bed in a dark, silent chamber for two 24-hr. periods, one week apart.

Procedure. The three groups, and the procedure before each session, were the same as in the original study except that: (a) Two new stimulus-scenes, similar in construction to the original two, were added, and (b) the *S*s were assigned to groups in such a way as to equate the three groups for average length of initial story.

---

* Received for publication October 30, 1964. This research was financed by Grant G-21762 from the National Science Foundation.

[1] Peter Suedfeld, Toward greater specificity in evaluating cognitive and attitudinal change, *Amer. Psychologist*, 19, 1964, 560; Peter Suedfeld, R. J. Grissom, and Jack Vernon, The effects of sensory deprivation and social isolation on the performance of an unstructured cognitive task, this JOURNAL, 77, 1964, 111-115.

[2] J. C. Pollard, Leonard Uhr, and C. W. Jackson, Studies in sensory deprivation, *Arch. gen. Psychiat.*, 8, 1963, 435-454.

After completing the first session, the *Ss* were rescheduled for another session one week later. In each case, the second treatment was identical to the first except that different stimulus-scenes were used.

Results. The initial stories of the three groups did not differ significantly from each other, in either the first or in the second session. The initial stories, averaged over all three groups, did not differ from the first to the second session.

Table I shows the average change in number of words per story from the initial to the final test in each session. In the first session, the *SD* and *SI* Groups showed a significant decrease and increase, respectively, while

TABLE I

CHANGES IN MEAN STORY-LENGTH
(Number of Words)

Treatment	First session		Second session	
	Mean	*t*	Mean	*t*
SD	−262.70	2.257*	−172.75	1.495
SI	+230.28	2.652†	+6.11	0.110
C	−56.27	0.961	−93.67	1.591

* P<0.02, one-tailed. † P<0.01, one-tailed.

the change in the *C* Group was not significant. None of the groups evidenced significant change during the second confinement period.

In Table I, the *N*s are 20 for the *SD* Group, 18 for the *SI* Group, and 15 for the *C* Group. In the *SI* Group, one *S* was discarded because the length of his initial story could not be matched. The data obtained from four other *Ss* from the *SI* Group are not included in the analysis because of a contaminating factor: during their first period of confinement, the ventilating system of the chamber was not operating, and they spent that 24-hr. session in temperatures of about 90°F. This environment was reported to be extremely unpleasant by all four of the *Ss*. For these four individuals, whom we shall call the *H*-Group, the initial confinement resulted in a mean decrement of 646 words; their second confinement, at normal room temperature, resulted in a mean increment of 161.5 words.

Discussion. The data obtained in the first experimental session confirmed our original findings. We again found that social isolation resulted in longer stories (interpreted as an attempt to maintain contact), and that the cognitive impairment which was caused by sensory deprivation counteracted this effect. Thus, the results again bear out the conclusion that, with the use of appropriate tests, introspective reports of sensory deprivation caused cognitive decrements can be verified objectively. The fact that these changes were not observed during the second

session indicates that both the motivational and the cognitive consequences of isolation and sensory deprivation can be negated through adaptation.

Data for the four *Ss* who underwent heat, tend to support our interpretation of the determinants of story-length. This *H*-Group, like the *SD*-Group, showed a consistent decrease in story-length after the first session. After the second confinement, on the other hand, these *Ss* behaved as the other *SI Ss* had when tested after the first session. The relatively severe unpleasantness experienced by the *H*-Group (all of whom spontaneously complained of the heat upon release) was associated with an average decrease in story-length which was approximately 2½ times as great as that found in the *SD*-Group; apparently, little if any adaptation to social isolation took place during this stressful experience.[3]

Besides making obvious the importance of controlling and specifying such conditions as temperature, the data imply that *SD* and *SI* may be situations along a broader stressful continuum. The consequences of such situations may be determined, or at least strongly influenced, by factors common to the entire range. Thus, we may be able to generalize from specific consideration of each single environment to a systematic consideration of a number of related environments.

Summary. Replication of a previous study showed that sensory deprivation adversely affects performance in an unstructured cognitive task. Sensorially deprived *Ss* told shorter stories after than before 24 hr. of confinement, while isolated but not deprived *Ss* told longer post-confinement stories. Both of these effects disappeared when the *Ss* underwent a second 24-hr. confinement.

---

[3] T. I. Myers, Studies of the stressfulness of sensory deprivation and confinement, *Amer. Psychologist*, 19, 1964, 740.

VISUAL-KINESTHETIC LOCALIZATION

By A. V. CHURCHILL, Defence Research Medical Laboratories, Toronto, Canada

During an experiment involving the lifting of unseen weights, it was observed that the *Ss*, when reaching for the weights presented directly in front of them, tended to move to the right of the weights with the left hand and to the left of the weights with the righ hand. It was noted also that they tended to look in the direction of the unseen object when reaching for it.

An experiment was designed to compare the accuracy of localization with the left and right hands when a position, located kinesthetically, was reported on a visual scale, and when a position, located visually, was indicated on a non-visual scale.

Apparatus. A visual scale was arranged horizontally and presented *S* with a concave 100° arc of a circle on an 18-in. radius, as shown in Fig. 1. The scale was graduated in 1° units and numbered in 10° steps from *S*'s left. A corresponding scale, not visible to *S*, was located immediately behind the visual scale. The apparatus was so arranged that, when the visual scale was presented at eye-level, *S* could raise his arm to shoulder height beneath the apparatus and, with his index finger, locate a stop (Condition 1), or when indicating a visually presented position, he could point to a position on the non-visual scale (Condition 2). A blackout cloth extended from just below the visual scale to cover *S*'s shoulders, completely eliminating any visual cues as to the position of his arm.

Subjects. Ten men, laboratory personnel, all reportedly right-handed, served individually under the two conditions. In both conditions, they were seated in front of the visual scale, facing its midpoint.

Condition 1. Under Condition 1, a stop was placed at one of a series of predetermined positions on the non-visual scale and *S* was instructed to locate the position of the stop with his index finger (in a manner similar to pointing a pistol). He was instructed to indicate, to the nearest degree on the visual scale, the position at which he was pointing. Nine different positions were chosen, and *S* was instructed to locate each position with the right or left hand, in either case moving from the extreme left or from the extreme right, and to report these positions on the visual scale. The 36 stimulus-conditions (positions × hands × directions) were presented in random order.

Condition 2. Under Condition 2, *S* was instructed to look at a specified position on the visual scale (the same nine positions used under Condition 1), and then

* Received for publication September 10, 1964. Defence Research Medical Laboratories Research Paper No. 572, PCC No. D77-94-20-23.

to indicate this position on the non-visual scale. Since this was a straightforward pointing procedure, *S* was not required to move his hand from either the extreme left or from the extreme right, as in locating the stops under Condition 1. The nine scale-positions were presented twice to each hand in a random order—a total of

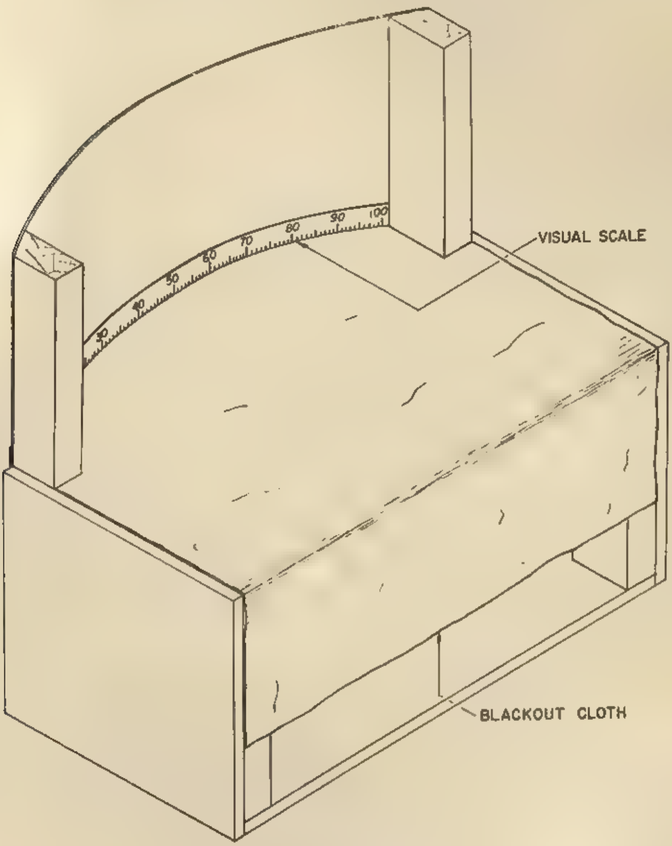

FIG. 1. FRONT VIEW OF APPARATUS

36 presentations. *S* was presented both conditions (Condition 1, then Condition 2) in a single experimental session.

Results. An analysis of variance showed significant differences between *Ss*, between conditions, and between hands ($p < 0.01$). The differences between directions of arm-movement under Condition 1 and between trials under Condition 2 were negligible. Responses to the different stimulus-

positions showed no significant differences. Hands × conditions and *Ss* × conditions were, however, significant ($p < 0.01$).

Hands × conditions is the result of the experimental design. If there is a bias in opposite directions with the two hands under Condition 1, and if the direction of the error is reversed under Condition 2, then this interaction must be significant. Thus kinesthetic overestimations to the right with the left hand, under Condition 1, would result in points located on the non-visual scale being called lower on the visual scale and, under Condition 2, the points located on the visual scale would be pointed out higher on the non-visual scale. Kinesthetic overestimation to the left with the right hand would produce the opposite effect.

The *Ss* performed in this manner under both conditions with the left hand, calling the points located on the non-visual scale lower on the visual scale under Condition 1 (Mean *CE* −3.07°), and pointing out positions higher on the non-visual scale than corresponding points located on the visual scale under Condition 2 (Mean *CE* +4.29°). If there is a corresponding bias with the right hand, it is apparently concealed by the greater accuracy of the right-hand responses (Condition 1, Mean *CE* −1.13°; Condition 2, Mean *CE* −0.29°).

Ss × conditions was a result of variability amongst the *Ss*. All *Ss* erred in the same direction, *i.e.* when the position on the non-visual scale was estimated on the visual scale, the *CE* was negative, and when the position on the visual scale was estimated on the non-visual scale, the *CE* was positive. The *S* with the greatest negative *CE* under Condition 1 also had the greatest positive *CE* under Condition 2, introducing enough variability to give statistical significance to this interaction.

As might be expected, the results obtained under Condition 2 logically follow from the results obtained under Condition 1, *i.e.*, if, under Condition 1, a stimulus on the non-visual scale were at 40° and *S* called it somewhere below 40 on the visual scale, then, when looking at the stimulus at 40° under Condition 2, *S* would tend to locate the point at a position somewhere above 40 on the non-visual scale. All 10 *Ss* performed in this manner.

Summary. The results of this study partly confirm the original observation, and further suggest that there is little difference in the accuracy with which *Ss* perform under the two experimental conditions, the main difference being in the directional bias of the error (mean *CE* −2.10° under Condition 1, and Mean *CE* +2.00° under Condition 2). The results suggest, also, that with right-handed *Ss*, performance is more accurate with the right hand (over-all *ME* 2.16°), than with the left hand (over-all *ME* 4.37°).

APPARATUS

CONTINUOUS MONITORING OF CALIBRATED TIDAL AIR DURING INSPIRATION AND EXPIRATION

By Roy J. Krusberg and Herbert Zimmer, University of Georgia

A combination of features deemed essential to a respiratory transducer for psychophysiological work, determined the performance characteristics of the transducer described here. This instrument was designed: (1) to provide calibrated measurement of tidal air flow; (2) to record the *S*'s respiratory activity without the artifacts introduced by loading his pulmonary system; (3) to yield separate and independent information on inspiration and expiration; (4) to permit continuous and prolonged monitoring over a period of many hours; and (5) to operate on direct current, and thus avoid the imposition of noise on other psychophysiologic transducers which are in simultaneous use. A.C. or carrier frequencies are likely to produce radiated interference in other transducers.

Instruments previously in use have not separated inspiration from expiration, and have recorded relative shifts in respiration-amplitude, rather than measuring tidal air flow in repeatable, standard physical units. A recent invention provides for the measurement of tidal air flow,[1] but it is: (a) apt to result in capacitive and inductive coupling of the *S* from the 200 k.c., 25-v. carrier frequency; (b) much more complicated and more critical in the tolerances required in the production and assembly of its parts; and (c), as a result, considerably more costly.

The transducers can be incorporated into any mask which covers the nose and mouth of human *S*s tightly, and contains separate valves for inspiration and expiration. For the present purpose, a Pulmosan Model C-263 respirator was modified by replacing the filter cartridge with a styrene tube 6 cm. long and 3 cm. inside diameter which holds the thermistor for inspiration, and attaching a similar styrene tubing to the exhaust valve to hold the expiratory sensor. The thermistor headers were mounted to place the thermistor beads in the center of the two tubes, 3

* This study was conducted in the Bioelectronic Computer Laboratory, of the University of Georgia, and supported by the Air Force Office of Scientific Research, Grant AF-AFOSR-257-65, and the Rome Air Development Center, Contract AF30(602)-3380.

[1] Hans Krobath and Corsan Reid, A new method for the continuous recording of the volume of inspiration and expiration under widely varying conditions, *Amer. J. med. Electron.*, 3, 1964, 105-109.

cm. from either end, and 1.5 cm. from the inside circumference. Coarse mesh screening was placed over the end of the tubes to prevent damage to the thermistors from curious fingers. A Brush 20-112 crystal microphone was installed 1 cm. proximal from the inspiratory valve to provide *S* with voice-communication.

The device used to sense the gas-velocity is a miniature thermistor, which is heated by current passing through a 10kΩ 1%-resistor that is placed in series with it and is provided with an OB2 gas diode voltage regulator.[3] The current flow through the thermistor is 8 m.amp., which heats the thermistor to about 200°C. Operation at this high temperature permits the use of an unbalanced input circuit, rather than a bridge-circuit, since the small temperature differential between ambient and exhaled air has a negligible effect on the resistance-level of the thermistor, as compared to the appreciable changes produced by the flow of gas. Thermistors have a negative temperature coefficient of resistance: as their temperature increases,

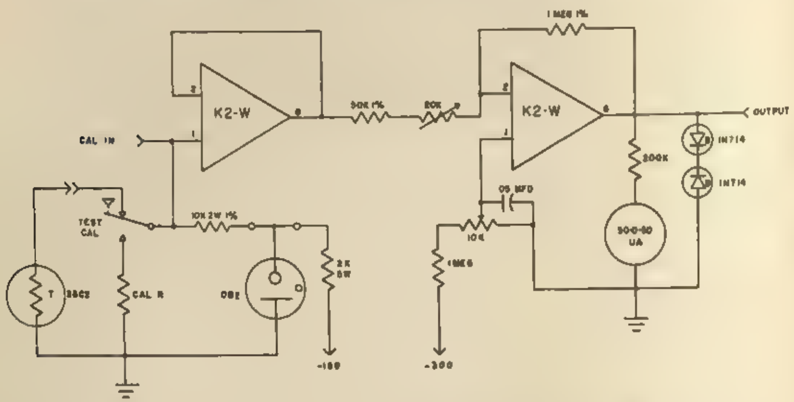

FIG. 1. DIAGRAM OF THE CIRCUITRY FOR THE RESPIRATORY TRANSDUCER

their electrical resistance decreases. At 200°C, the resistance of the 38C2 thermistor is 170Ω, or about 2% of its 8kΩ-rated resistance at 25°C. The resistor in series with the thermistor, together with the OB2-v. regulator, acts as a constant source of current, making the voltage across the thermistor a function of its temperature. When gas flows past the thermistor, heat is lost at a rate proportional to air velocity. As heat is carried away from the thermistor, its temperature drops and its resistance increases, resulting in a higher voltage drop across it. With the arrangement used in this transducer, the output voltage is a nearly linear function of gas velocity. Recovery time of this thermistor, from an instantaneous drop in air-flow from 0.5 to 0 l. per sec., is of the order of 100 m.sec., as estimated from an oscilloscope. The advantage of a thermistor over a thin, heated wire lies in its considerably greater sensitivity. A gas flow of one-half liter per sec. over the thermistor yields a change in voltage output of almost 1 v., whereas the voltage change for a hot wire, given this same gas velocity, would be in the order of m.v.

[3] Manufactured by the Victory Engineering Corporation, Springfield, N.J. The device is AXC496 (3812 mounted on a header.)

Fig. 1 illustrates one of the two identical control circuits employed. Two separate amplifier circuits and output indicating meters are provided, one for the inspiratory transducer, another for the expiratory transducer. To prevent loading the thermistor by active elements in the input circuit of the operational amplifier, it is necessary to connect the thermistor to the input of a buffer amplifier, with very high input impedance, in this case over 100 MΩ. The output of the buffer amplifier is fed to the input of a second Philbrick K2-W amplifier, with input and feedback resistors tailored to give the desired gain and zero point for a particular utilization.

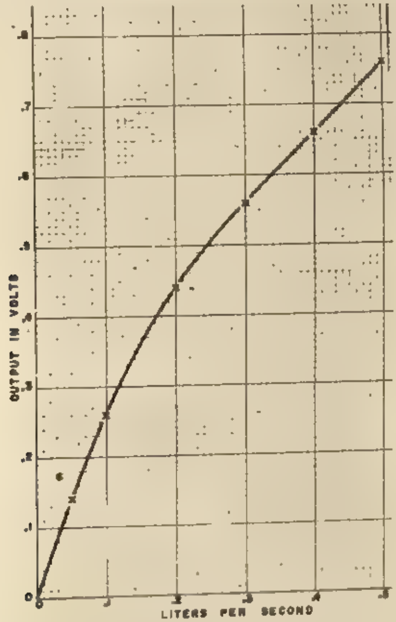

FIG. 2. AIR-FLOW CALIBRATED IN VOLTAGE EQUIVALENTS

In its present psychophysiologic application an output voltage range from −10 to +10v. is required to match the input of an analog-to-digital converter. The output amplifiers in this case are set for an amplification factor of 16, and a −10 v. zero offset. Two 10v. zeners are connected back-to-back across the output of each channel to prevent voltage outputs in excess of ±10v. from causing possible damage to associated equipment. Controls to set the amplifiers to zero are mounted on the front panel. A calibrate test-button inserts 220Ω fixed resistors, in place of the thermistor, for checking amplifier gain.

The thermistor gas velocity sensors were calibrated to 0.5 liters per sec. with a Fischer-Porter Model 10-A-3500 Flowrator variable-area flowmeter. The thermistors, mounted in their styrene tubes, were attached to a plenum, or pressure stabilizing chamber, which feeds the flowmeter. The output of the flowmeter was connected

by flexible rubber tubing to a venturi on the input of a shop vacuum cleaner. An adjustable pinchcock on the rubber tubing was used to control air flow through the system. While the accuracy of this respiratory transducer is better than $\pm 1\%$—a limitation is imposed by the calibrating instrument—the flowmeter, which has a full-scale accuracy of $\pm 2\%$. Fig. 2 shows the calibration curve of the transducer in terms of rate of flow in liters per sec.

In present use, the amplifier outputs are recorded in digital form on magnetic tape[3] once every 100 m.sec. for inspiration, and at an equal sampling rate for expiration, on two separate data channels. Tidal volume is calculated by summing the number of liters recorded at each m.sec. interval for any particular cycle of inspiration, expiration, or respiration. The duration of inspiration or expiration is determined to the nearest 100 m.sec. by calculating elapsed time between the -10 v. readings, which mark the beginning and the end of each cycle. If the output of the amplifier is connected to a graphic recorder, the tidal volume can be approximated by use of a planimeter.

---

[3] Herbert Zimmer, Preparing psychophysiologic analog information for the digital computer, *Behav. Sci.*, 6, 1961, 161-164.

NOTES AND DISCUSSIONS

THE GOLDEN SECTION REVISITED: A PERIMETRIC EXPLANATION

The golden section is a geometric proportion which has been of concern to philosophers and mathematicians since ancient times. Turnbull tentatively traces the concept of the golden section at least back to the building of the Great Pyramid in ancient Egypt.[1] Although the golden section has been mathematically defined in a number of fashions, one of the most popular definitions is that: $1:X = X:1\text{-}X$, or the extreme and mean ratio: the whole is to the larger part as the larger is to the smaller. This division is termed the golden section. To have the golden section in a rectangle, the width must be 0.618+ of the length.

Fechner found that people, when given a choice of a number of rectangles and asked to choose the most pleasing, actually selected most frequently one which possessed a ratio of 0.62 (ratio of width to length).[2] Actually, there was a rather broad preference mode in Fechner's data extending from a ratio of 0.57 to one of 0.67. Woodworth briefly surveyed the study of the golden section in psychological research and reported that "Rectangles have been tried by several experimenters [other than Fechner] with results similar in general to those of Fechner."[3] Some notable exceptions, however, do exist, *e.g.* Davis had students draw "most pleasing" rectangles and found few which possessed dimensions similar to that of the golden section.[4] Ogden has suggested that many experiments which had subjects (*S*s) "draw a satisfying figure, or to select from a number of figures the most favored" have not controlled a number of important experimental considerations which have influenced their results.[5] Woodworth argues against a simple ratio or mathematical explanation regarding the empirically found preferences for rectangles whose width are about 6/10 of their length, and concludes that "what the subject sees is not a ratio but a shape. In most cases, he is unaware of any specific ratio. If he thinks of the rectangle as a visiting card or as having

[1] H. W. Turnbull, The great mathematicians, in S. E. Newman (ed.), *The World of Mathematics*, 1, 1956, 80.

[2] R. S. Woodworth, *Experimental Psychology*, 1938, 384-385.

[3] Woodworth, *ibid.*, 387.

[4] F. C. Davis, Aesthetic proportion, this JOURNAL, 45, 1933, 298-302.

[5] R. M. Ogden, Naïve geometry in the psychology of art, this JOURNAL, 49, 1937, 210.

a particular use, his choice will be governed by that idea, but apart from use it would seem that any rectangle is potentially pleasing, the same as any color."[6]

Most psychologists will agree wholeheartedly with Woodworth's evaluation; however, it still does nothing to explain the reason(s) why so many

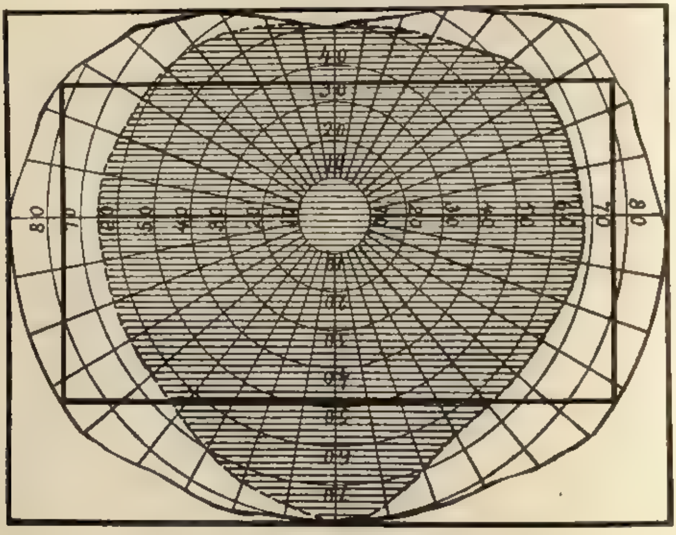

FIG. 1. THE BINOCULAR VISUAL FIELD FITTED WITH INTERNAL AND EXTERNAL RECTANGLES

*S*s, on at least two continents and over a number of decades, have shown extra preference for rectangular forms whose widths are approximately 6/10 of their lengths. The present note is an attempt to provide an hypothesis which we advance—and which we believe is testable—to explain this preference and the variability usually associated with it.

When a person's two eyes are fixed on a point ahead of him, a perimeter may be used to chart the visual field of each eye. There is considerable overlap in the visual fields of both eyes. The binocular visual field has been found to be of a somewhat irregular shape which is wider than it is high (Fig. 1).[7] With a little imagination, this binocular visual field can be seen to possess an outline form which is not too unlike that of a rectangle.

[6] Woodworth, *ibid.*, 388.

[7] T. C. Ruch, Binocular vision and central visual pathways, in T. C. Ruch and J. F. Fulton (eds.), *Medical Physiology and Biophysics*, 18th ed., 1960, 453. Permission to use a portion of the original figure has been granted by the W. B. Saunders Company and Dr. Ruch.

To pursue this line of thinking, the outline of the visual field was photographically enlarged and traced on graph paper. Interior and exterior rectangles (Fig. 1) were drawn within and outside of the field's outline form. An *average* rectangle, based on the averages of the dimensions of the internal and external rectangles, was computed, and the ratio of the height to the width of this *average* rectangle was found to be 0.665.[8]

By utilizing the metric of the graph paper employed (squares ¼-in. on a side), the area within the *average* rectangle was found to correspond well in excess of 90% (based on a very conservative estimate) to the area within the outline of the visual field. This rather good fit of the *average* rectangle to the visual field suggests that this would be an appropriate figure to use in describing the shape of the binocular visual field.

The fact that the binocular visual field may be regarded as being similar to a rectangle which possesses height and width which correspond somewhat closely to the properties of the golden section, provides the substance for a hypothesis which attempts to explain why rectangles having proportions similar to that of the golden section are generally regarded as having the most pleasing appearance. We do not wish to hypothesize *how* the dimensions of the common binocular visual field comes to be associated with esthetic judgments for rectangles, but it would not be difficult to construct a theory based either on *imprinting* or adaptation-level.

Dimensions of the visual field are partially dependent upon the size of certain facial features: the nose, eyebrows, and cheekbones. If there is a relationship between the outline of the binocular visual field and preference for form (rectangle and other geometric designs), one may expect that different combinations of facial features would have some effect upon the visual field and hence some influence upon rectangle preferences. Differences in facial features could then be expected to account for some of the variability that has been found to be associated with rectangle preferences. Conceivably, the observation made by Woodworth that American students seem to prefer more slender rectangles than those found in studies in France and Germany may possibly be related to differences in facial features.[9]

It is interesting to note that the height- and width-dimensions for many objects which are especially designed for visual presentation or appreciation bear a great deal of similarity with the dimensions of the binocular visual field. The height-width ratio for still and moving pictures projected

---

[8] The authors wish to acknowledge the collaboration of Mrs. Jeanette R. Collins in determining the dimensions of the *average* rectangle.

[9] Woodworth, *ibid.*, 387.

from 35 millimeter projectors is approximately 0.66. The average height-width ratio for a sample of 10 licenses, membership and business cards (randomly selected from a larger sample of such cards contained in the senior author's wallet) was found to be 0.653. We noticed that a great number of objects and devices which serve to limit visual fields, *i.e.* windows, picture frames, photographs, possess height-width or width-height ratios similar to that originally given as the golden section and for the binocular visual field.

Mental Health Research Institute — LeRoy A. Stone
Fort Steilacoom, Washington — L. Glenn Collins

THE CEILING OF PSYCHOPHYSICAL POWER FUNCTIONS

Sensory scaling by direct methods, such as fractionation or magnitude-estimation, usually yields a power function relating physical stimulus and subjective magnitude, at least among prothetic continua.[1] Most experiments, however, do not explore the upper end of the *O*'s dynamic range, presumably because *E*s are concerned for their equipment and their *O*s. Furthermore, the precise measurement of intense stimuli often encounters technical difficulties. Therefore the question of whether there are possible deviations from the power function with strong stimulation is as yet unanswered, though existing data seem to indicate that the power law is valid over the whole dynamic range. In any case, for loudness, the power law seems to endure stronger stimulation than do the earphones used in presenting the stimuli.[2]

An experiment in which subjective force was scaled by the method of magnitude-estimation provided an opportunity to investigate whether deviations from the power function were present in the psychophysical function when the stimulation was as intense as possible, *i.e.* near and at the upper limit of the *O*'s capacity. In this experiment *O* exerted the force by pushing a pedal in a horizontal forward direction with his right foot. By means of a beam scale connected to the pedal, *O* was presented with different stimulus-weights whose subjective magnitudes he estimated. The apparatus and procedure have been described in an earlier article.[3]

* This work was carried out at the Psycho-Acoustic Laboratory, Harvard University, and was supported by Grant G-10716 from the National Science Foundation. Report PPR-282.

[1] S. S. Stevens, On the psychophysical law, *Psychol. Rev.*, 65, 1957, 153-181.

[2] Stevens, The measurement of loudness. *J. acoust. Soc. Amer.*, 27, 1955, 825.

[3] Hannes Eisler, Subjective scale of force for a large muscle group, *J. exp. Psychol.*, 64, 1962, 253-257.

For each *O*, the set of stimuli contained a weight that exceeded the particular *O*'s force, and each *O* estimated the force he exerted in trying unsuccessfully to move the pedal. This procedure provided an upper end-point for the subjective scale of each *O*.

Four outcomes were conceivable: (1) *O* might refuse to give an estimate on the grounds, perhaps, that the subjective dimension had been changed from force to pain; (2) his estimate might be extremely high, perhaps even infinity, which would indicate a sharp rise in the psychophysical function at its upper end; (3) his estimate might coincide with or be only slightly higher than the estimate for the highest stimulus-force presented that fell below his maximal force, which would indicate an asymptotic leveling off of the psychophysical function at its upper end; or (4) his estimate could agree with the power function covering the lower part of the dynamic range. In fact, the *O*'s estimates agreed with the curve determined for the lower part of the range.

The stimuli, 10 in number, were spaced in equal log steps (0.204) between 10 and 630 lb. Ten men and 2 women each made 4 estimates of all the stimuli that fell below their maximal force and four estimates of the stimulus that fell immediately above it, which established the subjective magnitude of each *O*'s maximal force. All the *O*s were able to use the lowest six stimuli, and seven of them could use the lower nine stimuli, but none could move 630 lb.

Treating the data in the customary way would have obscured any deviations in the upper part of the power function, because the maximal force for the different *O*s was distributed along almost half the range of the function. Therefore the following procedure was substituted.

An *O*'s maximal force was regarded as a sort of upper threshold or "ceiling" and assigned a decilog value of zero. In the case at hand, the exact ceiling was not determined since it did not seem probable that different *O*s would tax their resources of force equally. Instead, the first stimulus-force above the particular *O*'s maximal force, *i.e.* the stimulus 2.04 decilogs higher than the strongest stimulus at which *O* succeeded was taken as his ceiling. The stimulus-forces could then be counted downward from the ceiling in decilog steps, by a method similar to the counting of decibel steps of sensation-level in loudness.[4]

A similar procedure was employed for the subjective scale. Since, according to Stevens' psychophysical law, equal stimulus-ratios yield equal subjective ratios, the decilog step used between adjacent stimuli should correspond to a constant decilog step between adjacent subjective magnitudes.[5] Thus, each *O*'s estimate of his maximal force was assigned zero decilogs, and his estimates for the other stimuli were transformed correspondingly into negative decilog values. The arithmetic means of these decilog values for all the *O*s and all the presentations was computed for each stimulus-step. Thus the estimates averaged together are not for the same stimulus in pounds, but rather for a stimulus representing the same number of physical steps below each *O*'s ceiling. If we imagine each *O*'s psychophysical function as a

[4] The use of decilogs for physical units in continua other than the intensity of sound has been recommended by Stevens. See Stevens, Decibels of light and sound, *Physics Today*, 8, 1955, 12-17.

[5] Stevens, *Psychol. Rev., op. cit.*, 162.

straight line in a log-log-plot, the procedure outlined above corresponds to a translation of the lines in the x- and y-directions so that the upper end points coincide. In essence, this is an individual scale transformation of both axes, obtained by multiplying by a constant.

Fig. 1 shows a plot of subjective magnitude in decilogs against force in decilogs. The power function is valid over the whole dynamic range, though the deviation

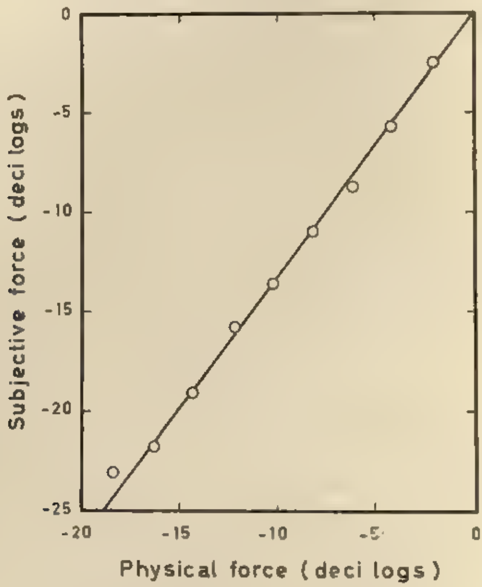

FIG. 1. SUBJECTIVE FORCE AS A FUNCTION OF PHYSICAL FORCE, BOTH IN DECILOGS DOWNWARD FROM THE *O*S' CEILING.

at the lowest point suggests the need of a correction for the lower threshold. When the same empirical data were treated in the customary way (described elsewhere), the deviations at the lower end were more conspicuous.[6] In Fig. 1 the deviations are obscured, since the values for the lowest stimuli for different *O*s distribute themselves along a large part of the range. This displacement of the absolute threshold is probably also at the root of the different exponents obtained —1.3 for the data treated as in Fig. 1, and 1.5 for the customary treatment of the data. Since the concept of a ceiling is somewhat questionable from the point of view of sensory scaling, 1.5 may be a closer approximation to the exponent for subjective force given by this experiment. The method employed in the present study was devised only to investigate the upper part of the psychophysical function.

Summary. Subjective force exerted by pushing a pedal with the foot

[6] Eisler, *op. cit.*, 254.

was scaled by the method of magnitude estimation. The validity of the psychophysical power law was demonstrated for the whole dynamic range up to the strongest forces the *O*s were capable of exerting.

University of Stockholm HANNES EISLER

AUTOKINESIS WITH AN AFTER-IMAGE

In an investigation of autokinesis, three viewing conditions were employed: (1) steady fixation of a small, low-intensity white light for 1.5 min.; (2) steady fixation of the light for 2.5 min. with the retinal image of the light initially adjacent the intense after-image of an arrow. The arrow was horizontally disposed, its tip pointing toward the light, from which it was separated by 49′ of arc for 19 *O*s and 72′ of arc for 6 *O*s. The maximal angular subtense of the arrow was 47′ of arc; and (3) observation of the after-image of the arrow, arranged to occur as near to the fovea as possible, for 1.5 min. The after-image was produced by placing an arrow-shaped aperture before a 'Mecablitz 502' electronic photographic flash-unit of rated flash-duration and power (0.001 sec. and 120 joules, respectively). A well-defined after-image of the arrow produced by this method remained visible for some 4 min. Twenty-five *O*s, 17 of whom had no previous experience of autokinesis, were exposed to the stimuli once under each viewing condition. The stimuli were viewed monocularly in a completely darkened room. The *O*s were provided with a head-rest and wore correcting spectacles if normally required.[1]

Under Condition 1, all *O*s experienced movement of the light. Under Condition 2, 19 *O*s reported movement of only one stimulus at any time, 15 of them reporting movement only of the after-image. Three *O*s reported movement of both stimuli at the same time, with varying separation between the light and after-image. They also reported movement of both stimuli individually. Three *O*s only reported movement of both stimuli together with unvarying separation between the light and after-image. Introduction of a second structure, the after-image of the arrow, appeared to reduce the reports of autokinesis. Under Condition 3, all *O*s reported movement of the after-image. Six *O*s reported that it moved completely out of their fields of view. Observed movement under all con-

[1] This work, on which this Note was based, was done at Stow College in Glasgow. The author is indebted to Mr. T. S. Moore for providing the laboratory space. The author wishes to thank, for their advice, Dr. C. R. Evans, National Physical Laboratory, Mr. M. Green, University of Strathclyde, and Dr. G. E. McKinnon, University of Waterloo.

ditions was primarily of a slow, drifting nature, along gently curved arcs. The velocity was reported to be similar to that experienced with *muscae volitantes*.

Movement of the after-image with respect to the light would appear to support an eye-movement explanation of autokinesis, but the reports of the after-image and light moving together with unvarying separation, reported also by Carr,[2] might support explanations involving central processes such as suggested by Gregory and Zangwill[3] and by Verheijen and Oosting.[4] The disappearance of the after-image from the field of view, reported by 6 *Os*, might have been a consequence of 'tracking' or 'chasing' an image slightly displaced from the fovea.

University of Waterloo DAVID J. PIGGINS

[2] H. A. Carr, *An Introduction to Space Perception*, 1935, 289-335.
[3] R. L. Gregory and O. L. Zangwill, The origin of the autokinetic effect, *Quart. J. exp. Psychol.*, 15, 1963, 252-261.
[4] F. J. Verheijen and H. Oosting, Mechanism of autokinesis, *Nature*, 202, 1964. 979-981.

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

The Behavioral Basis of Perception. By JAMES G. TAYLOR, with a mathematical appendix to Chapter 8 by SEYMOUR PAPERT. New Haven and London, Yale University Press, 1962. Pp. viii, 379. $10.00

Taylor's seductive thesis is that *all* conscious experience is the result of acquired specific readinesses to respond to environmental objects, a remarkably thorough-going empiricist position which, while it is not new, has not yet, in his opinion, been subjected to critical tests because it had not previously been stated with sufficient precision to allow specific predictions about the extremely complex system of man-in-environment (with its multidimensional independent variables) to be calculated and verified, nor has it been traced from the beginning of life through the stages by which consciousness is built up. It is to these two tasks that this book is directed.

At some point during the fifteen years it has been in preparation, Taylor deliberately renounced reading current literature which only interfered with his construction of a consistent theory. His basic debts are these: To Ashby, for the general concept of the multistable system (here, the organism in its environment), and to Hull, for the more specific formulations of how response tendencies are established and of how responses are elicited by the afferent neural consequences of sensory stimulation. The basic model is as follows.

The developing human organism in its environment comprises a system with so many variables that, if it should enter a *critical state* (one which threatens to disrupt the system), those variables could not reasonably be expected to take on simultaneously, by random changes alone, just that combination of values which would adapt the system (*i.e.*, would move it out of the critical state). As a *multistable system*, however, it contains many smaller subsystems, each of which, with fewer variables, can adapt (serially or in parallel) in isolation. Adaptation occurs when a system (or subsystem) is thrown into a critical state. The system then randomly adopts different values of each of its variables, until a combination results which removes it from that state (in Hull's terms, S_D in conjunction with an external stimulus S_x generates unlearned behavior until some response diminishes the drive) with retention of those values that end the critical state; retention is due to neural connections or *engrams*, formed in the process of conditioning. Initially, in an infant that can move only its head, eyes and arms (leg movements being completely ineffective as far as looking and reaching are concerned), the system needs only to include the afferent variables necessary to produce a response of reaching (from any of the possible positions of the hand and eye) for an object at any point in space. Here the drive states involved are those that are reduced by grasping an object which can be carried to the mouth or which provides support against falling movements. Invoking Hullian principles (*e.g.*, in generalization $_SH_R$ is a decreasing exponential function of the differences in jnd units between stimuli), Taylor

shows with great care (using set-notation, with a section provided to instruct the reader in its use) that equivalence classes of afferents (visual, kinesthetic and drive) could be built up, each member conditioned to reaching responses which terminate at the same object at some point in space, from any orientation of head and eye, and with any initial position of the hand. Engrams (neural traces) mediate those equivalence classes. The engrams that are energized (by drive), that are active at any moment, and in a state of readiness for the execution of each response in the set (only one of which can be executed at any moment), Taylor identifies as *conscious perception*. No matter what the orientation of head, eye and hand may be, the organism is always prepared to respond to any fixed object in space by a movement which will bring the hand in contact with the object. In consequence, the positions of objects in space are perceptually invariant under the different orientations and movements of head and eye; the first of our perceptual constancies has been established, and these constancies themselves provide stable bases on which to build more easily all of our further adaptations.

In progressive stages, each of which builds on the previous stages, as the infant learns to sit, stand, and finally walk, successive adaptations occur to more and more variables. With the ability to sit up and move its trunk, a perceptual frame of reference is built up and added to all of the properties of the previously adapted subsystem; this enables the perceiver to refer the position of any object in the environment to all other objects' positions, but it does not yet include differential responses to objects' distances beyond arms' reach. With the ability to stand, stable horizontal-vertical axes are built into the framework (and both the general results and the patterns of individual differences found by Witkin and his colleagues are explained). With the addition of locomotion, the movement-dependent depth cues provide new equivalence-classes; the radius of the perceived world is enlarged. Since, regardless of the course of distortions undergone by an object's retinal image, the final handling response remains invariant in certain respects, the classical constancies (size, shape, etc.) are acquired. Perceived space at first approximates Euclidean properties only within a very narrow range, which expands with the child's capacities for dealing with the world, but which is never complete.

The basis of the apparent immediacy of sensory experience and, specifically, *i.e.* the origins of the different modalities of sensory (and affective) experience are to be explained completely in terms of the kinds of responses to which different classes of stimuli have become conditioned: *e.g.* an auditory stimulus source behind the head evokes a turning response impossible with visual stimuli; taste and smell differ because only olfactory stimuli can become conditioned to the response of putting the emitting object in the mouth, and are allied because they jointly evoke subsequent responses (the associated interoceptive responses provide a pleasant-unpleasant component, depending on whether they are, respectively, evoked or inhibited, *e.g.* by full stomach, etc.); emotions are distinguished from each other by the engrams mediating the overt responses to which each is conditioned.

The primary opponent is *nativism*, a portmanteau here meaning any view other than one in which *all* consciousness (including so-called raw or primitive sensory experiences) is acquired through learning. Gibson is the chief target, but Hebb's theory is also dismissed as self-contradictory in its axiomatic bases, and even Piaget, despite his heavy emphasis on ontogenesis, is revealed as positing an

untestable, primitive consciousness—unparsimonious because it postulates both an innate mechanism and an acquired mechanism to account for the same phenomenon (consciousness). The closest predecessor of Taylor's viewpoint is Hayek's (considered in a final chapter on "Philosophical Implication"). To Hayek, as to Taylor, a primary impulse from the receptors has no significance in itself; only its "acquired following"—the subsequent impulses it generates—gives it sensory quality. Hayek, however, provides no mechanism which determines what these followings shall be, an oversight that leads indirectly to his deduction that the brain cannot classify its own classifying acts, making it impossible to explain the mental order (consciousness in toto) in terms of physical events—which is just what Taylor essays. To Taylor, we have seen, a primary impulse from a sense organ means nothing except in terms of its membership in a conditioned response system: each sense datum is in itself an act of knowing—"the realization of a potential relation between the organism and a detail of the environment," and the simultaneous knowing of a large number of facts about the environment is what is meant by being conscious. This is by its nature private knowledge (being each organism's own field of responses), which can be studied only indirectly—but which certainly can and should be studied, behaviorist though Taylor is.

In criticism, it is easy to quarrel with Taylor about details of content. Although the book is written with vitality and great care, it is unfortunately casual in places. For example: Gibson is denounced for asserting, that the unbounded "visual world" is based on the "visual field" (which is, in contradistinction, limited to a small region around the fixation-point); in fact, Gibson has repeatedly denied the priority of the visual field, considering it derived, artificial and uninformative with respect to the problems of perception of the visual world. Again: Taylor never seriously tackles the precise nature of the relationship between the various kinds of perceptual task, *i.e.* introspection, search, incidental perusal, preparation to act (incidentally, Brentano and Brunswik deserve footnotes at least in any such theory)—problems that would seem to me to be central to any attempt to identify perception with specific acquired response-system readinesses. As a third example, I can't see why Taylor is willing to concede the possibility of two mechanisms to subserve "color contrast," *i.e.* his operator which adds the complementary of the surround to each object's perceived color (a statement which seems to me not one whit more behavioristic than Helmholtz' earlier version) and some mechanism of retinal irradiation (at least in principle, since he then adduces reasons for rejecting this latter mechanism), yet finds the idea of two mechanisms underlying consciousness (say, the ontogenetically prewired engrams subserving primitive experience and the acquired engrams subserving conditioned response systems) so unacceptable.

There are other points about which one might quibble, of course, but there is a more serious point of criticism to raise.

Taylor argues that theories must be rigorously thought through before any research is undertaken (Erismann and Köhler are chided for rushing too soon to experiment), and the book maintains an air of sequential, successive and rigorous deduction. This is misleading, however, since Taylor has revised his thinking to fit the findings that he does consider (pp. 211 and 268), and these findings form a very highly selected group of data. This marks what is weakest about the book:

it both profits and suffers throughout by being too simple in its summary of data: perceptual change is not necessarily dependent on any overt response, and neither the facts of adaptation nor the "achievements" of the perceptual constancies are as complete and unequivocal as casual reading of his arguments would suggest. Moreover, Taylor repeatedly presents as an unavoidable consequence some prediction for which it seems easy to arrange alternative options: his own system is less vulnerable than he takes it to be (and so, I believe, are those he demolishes)—either that, or it's false.

Nevertheless, the attempt at so thoroughgoing an empiricism certainly makes this book an important one, and it is packed with ingenious explanations, embedded in a seductive and thought-provoking enterprise.

Cornell University JULIAN HOCHBERG

The Approval Motive: Studies in Evaluative Dependence. By DOUGLAS P. CROWNE and DAVID MARLOWE. New York, John Wiley and Sons, 1964. Pp. xii, 233. $7.50.

The authors of this monograph both received their Ph.Ds. in 1959. Shortly thereafter they collaborated in developing a True-False personality scale modeled after the Lie scale of the MMPI. Most of the 18 True keyed items in the Marlowe-Crowne (M-C) scale include the modifiers *always* or *never*. An example: "No matter who I'm talking to, I'm always a good listener." Most of the False keyed items include the modifiers *at times, on occasion,* or *sometimes*. An example: "I like to gossip at times." If subjects are aware of the modifiers in these two statements, then it seems reasonable to believe that relatively few of them could truthfully answer the first item True and the second item False. The same considerations apply to the other items in the M-C scale. Yet the mean score on the M-C scale for college students is typically about 15.

The main body of the monograph is a report of a series of studies primarily done by Marlowe, Crowne, and their graduate and undergraduate assistants in which the results are interpreted as supporting their interpretation of high scores on the M-C scale as indicating need for social approval. The typical research design used in these studies involves testing subjects under the same or varying conditions. The subjects are then divided into high and low groups on the basis of their scores on the M-C scale. Differences between the means of the high and low groups on a dependent variable, when found, are then interpreted as resulting from differences between the groups in need for social approval.

Numerous and various tests were used and the results, in general, turn out to be in accordance with the expectancies of the authors. Thus they expect high scorers compared with low scorers: (1) to describe a dull and repetitive task as more enjoyable and of greater scientific importance; (2) to show greater conditionability in response to an experimenter's "Mm-hmm"; (3) to show greater agreement with the "inaccurate" number of knocks that confederates of the experimenter report they have heard; (4) to regard more favorably an accomplice who has, in fact, tricked them; and so on. It can also be argued however, that low scorers compared with high scorers are more honest and truthful. The truthful subject might be expected to describe a dull and repetitive task unfavorably; to report the actual

number of knocks he has heard; to express how he feels about being tricked, and so on. There is no compelling evidence reported in any of these experiments that establishes that high scorers have a strong need for approval and low scorers do not Some low scorers may also need approval, but, at the same time, they may be more truthful than high scorers.

Crowne and Marlowe are aware of the alternative explanation of these experiments, but they believe that the results obtained in the verbal conditioning experiments cannot be interpreted in terms of what they describe as the "liar" hypothesis. Let us consider, therefore, these experiments in greater detail.

There is evidence, from other studies, to indicate that verbal conditioning occurs only when subjects are aware of the reinforcement. In the studies reported in this monograph, the authors claim that "aware" subjects were eliminated. But were they? Consider the fact that they find no evidence of conditioning for the two controls and the low experimental group and that this result is consistent with other studies which fail to find conditioning for unaware subjects. The high and, presumably, "unaware" group, on the other hand, does show an increase in plural noun responding. Could it be that at least some of these high scorers were, in fact, aware, but failed to report it? If high scorers on the M-C scale are "liars," then perhaps some of them also lied about being "aware" of the nature of the experiment. This study does not provide a clear refutation of the "liar" hypothesis. In fact, in a study *not* carried out by the authors or their associates, Spielberger, Berger, and Howard found that need for approval, as measured by the M-C scale, was not related to the subjects' reinforcement motivation or performance. Only those subjects aware of the correct response-reinforcement contingency showed a significant change in performance over trials and this was true for both high and low groups, the unaware high and low groups showing no significant change.

In a second verbal conditioning study, reported by the authors, subjects were reinforced with "Mm-hmm" for positive self-references. It is stated that 34 men and 42 women were assigned at random to either the experimental or control groups. It is also stated that these subjects were questioned to determine the extent of their awareness, but, presumably, *none* was aware because the results are reported for the total of 76 subjects. That not one out of 76 subjects is considered by the experimenter to be aware seems unusual. This is undoubtedly a lower rate, if it is a rate at all, than that found by other investigators. The reviewer is left with the completely subjective impression that either the interrogation was not thorough or that some subjects may possibly have lied when they were interviewed about the nature of the experiment. In this experiment subjects were again assigned to high and low groups on the basis of their scores on the M-C scale, but, in this instance, the high experimental group has 24 subjects and the low group only 14. Ordinarily, in the other studies, the division of the subjects at the mean results in groups which are fairly comparable in numbers. No explanation is given for the discrepancy in this study. Nor is any information provided as to the number of men and women in each of the groups. Would it make a difference if there were an excess of men or women in the various groups?

The authors state that there is a significant increase in the mean number of positive self-references from the first to the second period for the high experimental group and this is so, the two means being 0.181 and 0.233. The mean of 0.233

for the high experimental group in the second period is almost exactly the same as the first period mean, 0.229, for the high control group. Thus, after 10 minutes of reinforcement, the rate for the high experimental group is approximately the same as the initial rate for the high control group. Would one really want to conclude, as the authors do, that the experiment demonstrates the effectiveness of reinforcement on the conditionability of high scorers on the M-C scale?

A point of some concern in several of the studies is the discarding of subjects without adequate explanation. In one verbal conditioning study a total of 145 subjects were randomly assigned to one of three conditions. Before the data were analyzed, 19 subjects were discarded, 15 of the 19 subjects being those assigned to the negative reinforcement condition because they were aware of the conditioning procedure. Presumably ⅓ or approximately 48 subjects had been randomly assigned to this condition initially. If 15 of these subjects were discarded *before* the data were analyzed but on the basis of the *post-experimental* interview, then there should be only approximately 33 subjects left in the negative reinforcement group. Yet, the authors state: "The final experimental sample of 126 consisted of 42 subjects in each of the three conditions" (p. 51).

In another experiment on perceptual defense 14 subjects were discarded because their responses in a post-experimental interview could not be reliably classified. Another 18 subjects were dropped chronologically in order to obtain equal *n*'s for the analysis of variance. Having reported the analysis of variance, the investigators proceed to base their major conclusion upon 12 *t*-tests. Since the *t*-tests do not require equal *n*s, was there really any need to discard the 18 subjects?

In a word-association study, 134 subjects took the test under relaxed and then under popular instructions. Instead of dividing the subjects into high and low groups at the mean on the M-C scale, as in the various other studies, the data are analyzed only for approximately the upper and lower 25 per cent of the sample. In the *t*-test for the relaxed instructions, the *n*s for the high and low groups are 35 and 41, respectively. Presumably, these *n*s would be the same for the *t*-test for popular instructions. Instead, 4 subjects have disappeared from the high group and 8 from the low group. No explanation is provided the reader.

The consistency with which the results of the various studies are in accord with the experimenters' expectations, reminds one of Rosenthal's work on experimenter bias. In fact, the authors state: "In several investigations, significant correlations were found between the need for approval and experimenter bias. The average correlation over five samples was 0.74" (p. 83). The authors believe that the studies reported in the monograph are free from this potential source of bias because the M-C scale was not scored for the subjects until after the studies were completed.

The experiments reported by Crowne and Marlowe are interesting and the variety of situations investigated is impressive. As is often the case in psychological research, however, the studies need to be repeated by other investigators.

University of Washington ALLEN L. EDWARDS

Risk Taking: A Study in Cognition and Personality. By NATHAN KOGAN and MICHAEL A. WALLACH. New York, Holt, Rinehart and Winston, 1964. Pp. x, 278. $8.00.

Within recent years, we have witnessed a growing dissatisfaction with the psychological application of econometric theories of decision-making. It has been

argued, and in my view correctly, that such application neglects the role of personality factors in risk-taking behavior. There has been scant empirical work, however, designed to demonstrate the relationship between personality and risky decision-making. The present book represents a systematic attempt to explore this area from the point of view of contemporary psychology. In a commendable effort at comprehensiveness, the authors made use of six different decision-making procedures, three cognitive-judgmental variables, and nine measures of personality-dispositions, all of which were analyzed separately for samples of men and women.

The book first presents data relevant to the generality of risk-taking behavior across hypothetical and pay-off contexts. Among the decision-making procedures used in the study was a Choice Dilemmas instrument consisting of hypothetical decision situations involving risky and safe courses of action. An *S*'s selection of the probability-level for the risky alternative's success that would make it just attractive enough to be chosen was taken to reflect a conservative or risky strategy. In the pay-off domain, several procedures were used, such as presenting the *S*s with pairs of bets on dice which varied in terms of probabilities of winning or losing, and the amounts of money to be won or lost. *S* chose the bet he would prefer in each of 66 pairs of bets. Conservative or risky strategy indices were derived from the pattern of an *S*'s selections. A second pay-off procedure yielded similar indices, but under skill rather than chance conditions. *S* played a shuffleboard game and made bets on his chances of shooting a puck between two wooden posts.

The evidence for an association between decision-making strategies under hypothetical and pay-off conditions was strongly positive for women and moderately positive for men. This suggests that contrary to some prevalent conceptions in the literature, the use of hypothetical decisions may in fact approximate more ego-involving choice situations. The findings, however, become more complex when subjected to moderator (*i.e.* internal) analysis, a technique developed by the authors in connection with their earlier work. Test-anxiety and defensiveness (as measured by the Alpert-Haber and Marlowe-Crowne scales, respectively) were chosen as moderator variables because of their apparent relevance to risk-taking behavior. These variables were used as moderators throughout the entire study in an effort to discover relationships among cognitive and decision-indices which might otherwise be obscured in correlations based on the total sample.

One of the more important issues in the area of decision-making is the relationship between decision-outcomes and postdecision-satisfaction. The authors present considerable data on this issue. For example, they report a significant tendency for low defensive-low test-anxious men (in a skill context) to show a degree of postdecision-satisfaction proportional to the amount of money they had won. For those high on both test-anxiety and defensiveness, the opposite behavior occurred; the less these persons won, the more satisfied they were with their bets. It would appear that failure made them even more committed to their particular decision, a result which takes on added significance when compared with recent dissonance theory-derivations concerning expectancy confirmation and disconfirmation.

The authors state that one of the main goals in their study was to examine the possibility that judgmental processes in cognition can be conceptualized in terms of making decision. Previous research has produced inconclusive results regarding the relationship between cognitive indices of judgments and measures of risk-taking. The authors suggest that past failures may have been due to the use of exclusively

hypothetical procedures in making decision in which *Ss'* ego-involvement was quite weak. The present study attempted to overcome this limitation by examining relations between the *Ss'* strategies in decisions under pay-off conditions and their cognitive styles in judgmental tasks, with the moderator variables of test-anxiety and defensiveness taken into account. Three judgmental tasks of cognition were used: (1) Pettigrew's Category Width Test; (2) Brim's Judgment Extremity-Confidence Procedure; and (3) the extremity of an *S*'s self-ratings of his personality. The results indicated that empirical relations do exist between cognitive tasks and strategies of risk-taking. The relations, however, were often complex, with pronounced sex differences almost always present. Thus significant associations between the domain of cognition and decision-making tend to be confined to women. The nonsigificant results for men change, it is true, when moderator effects are taken into account, but the direction of these effects is also sex-linked.

The remainder of the book is devoted to examining the influence of intellective abilities on decision-making, and the association between personality dispositions (other than test-anxiety and defensiveness) and decision-making variables. The general trend of the results suggests with some justification the novel idea that verbal ability itself may be partially determined by risk-taking dispositions rather than the reverse. Analytic vs. global forms of intellective functioning were also related to behavior in risk-taking with sex differences once again exerting a strong influence.

Hypotheses regarding direct personality correlates of risk-taking received weak confirmation. Personality dispositions such as self sufficiency, independence, impulsiveness, and rigidity showed only low level associations with decision-strategies for samples of men and women. Some effects were observed, however, when the moderator variables were introduced. It would seem that when dealing with a particular personality variable, the investigator must pay careful attention to other personality dispositions that operate as a motivational context.

It cannot be denied that this book is an impressive and illuminating piece of work, but its very comprehensiveness leads to an unfortunate eclecticism. Instead of following a limited theoretical approach from which specific hypotheses were derived and tested, the authors included a great many variables simply because they appeared relevant to each other on general theoretical grounds. The result is a vast amount of data which often makes interpretation rather difficult. Then, too, we are faced with large numbers of correlations for which *ad hoc* explanations are introduced, ultimately preventing the authors from formulating a comprehensive conceptual scheme. In addition, such a large number of correlations may yield a certain number that are significant solely for reasons of random fluctuation. Which of the many coefficients are due to random effects is of course, difficult to specify. The absence of a theoretical framework which the reviewer had hoped to find in a study of this proportion may be largely due to the nature of correlational analysis. Even with the use of moderator-variable analysis, correlational studies simply do not establish the causal relationships necessary for integrated psychological explanation.

The present book inevitably invites comparison with previous work in the area of personality and decision-making. I believe that Kogan and Wallach have here contributed to the best of research in this field. The book should help to bring the behavior of risk-taking into line with the main body of psychological thinking

and free it from the confines of static econometric theory. If only for this reason, the book deserves a careful reading by psychologists and others interested in personality and risk-taking.

Russell Sage Foundation DAVID C. GLASS

A History of Psychology by ERWIN A. ESPER. Philadelphia: W. B. Saunders Company, 1964. Pp. vii, 368. $6.50.

In this fairly short book, Professor Esper says he wrote not *the* history, but *a* history, of psychology. To explain his purpose in writing a fairly selective book, he starts from the thesis that psychology should be (and once was, under Aristotle), the study of man as a biological organism. This proposition then provides the basis both for a detailed treatment of the naturalistic psychology, as a branch of biology, of the Greeks, and for repeated attacks on those movements which, in the author's view, have distorted psychology away from its primary goal. Such movements include centuries of religious dogmatism, the mentalism of Wundt, the applied arts of social and clinical psychology, and psychoanalysis, the empty-organism approach of some behaviorists, and the complex of phenomenalism-Gestalt-existentialism. Thus, the book is not simply descriptive; the author's biases are intended to show.

After an introductory chapter on reasons for studying history, four chapters are devoted to the origins of magic, animism and religion, the development of the naturalistic approach by a number of Greek philosopher-physicists, the more sophisticated mysticism and anti-naturalism of Pythagoras, Socrates, and Plato, and to the beginnings of biology and medicine. Throughout these and later chapters, the author occasionally traces the continuity of an idea from Greek to modern times, as in his relating of the doctrine of pneuma, from the Greeks and Galen, to the hydraulic theory of Déscartes, to neural drainage explanations of learning, to libido theory, and finally to the hypothesis of growth of synaptic knobs.

The next three chapters, nearly one-fourth of the book, deal with Aristotle: his logic and his linguistics of science, his view of man as a biological organism, and his view of man in society. Esper's summary of Aristotle's linguistics enables him to elaborate on another theme that recurs throughout the book, viz. that some of psychology's problems, such as the mind-body distinction, are only linguistic tangles and should be seen as such. The chapter on biology, which includes psychology, is an important one for the author since he seems to hold that under Aristotle, psychology reached a kind of peak, in some ways never reattained, as an empirical science devoted to the study of the individual organism as a biological unit. The third chapter deals with Aristotle's development of a social psychology, based, the author says, in part on several Aristotelian principles of individual behavior. Thus, at least with Aristotle, social psychology was built on the basic science of individual behavior.

The eight chapters considered so far make up about two-thirds of the book. Three chapters on psychology as philosophy, as social science, and as a biological science, conclude the book, except for references and an index. These three chapters are rather heterogeneous; each gives the impression that it could, and probably should, have been expanded into several separate chapters. Each includes historical sketches, usually brief, of one or more topics from Aristotle to the

present. It is also in these chapters that the author's strong views on what psychology should be, in contrast to the grab bag of activities that it presently involves, are most evident.

Esper's drubbing of philosophers, particularly those latter-day oracles, the philosophers of science, should delight all die-hard empiricists. Deplorably, even positivism now appears to be returning to metaphysics. In reviewing the seemingly-endless parade of verbalisms and irrationalisms, from ancient philosophy to existentialism, that continue to burden psychology, he downgrades, as mentalists, Wundt and Fechner. In contrast, he sees Helmholtz and Ebbinghaus as founders of a true science of psychology. This view of these two pairs of men is interesting and should somewhere be presented in more detail; it also suggests that history should be written by historians with different points of view.

The chapter on psychology as social science is chiefly devoted to analysis of the work of McDougall and Weiss. McDougall's antimechanistic social psychology is lambasted. Esper thinks that both McDougall and Freud, like Wundt and Fechner, have been vastly overrated. Weiss' extreme mechanism fares considerably better, though he finds little in Weiss that was not already said by Aristotle. He concludes that a biologically-oriented psychology should not really be grouped with social and clinical psychology, or with the other social sciences. He predicts that the term "psychology" will continue to be used by practitioners, but holds that it has lost its usefulness both for science and as a university administrative category.

The final chapter on psychology as a biological science covers the history of neurophysiology rather extensively for a book of this size. The reason for this emphasis seems clear: Esper flatly maintains that the basic science of man really is or should be the science of physiology. Those whose interests are in the development of a basic science of man, whether they call themselves physiologists or psychologists, should be in the same university department. Mentalistic psychologies, and psychology, as a collection of applied arts, would, of course, not be included. In propounding this view, he also sharply criticizes the peripheralist and empty-organism types (Loeb, Watson, Skinner), as well as the pseudoneurology of hypothetical constructs and intervening variables.

In Esper's book, there is more history of psychology than one might suspect. Many men and ideas are mentioned, though often very briefly. With regard to this, the historical aspect of the book, one can say that it is a fairly good, short, history of psychology. The author's polemics, directed against essentially every variety of psychology that is aphysiological, are, however more difficult to evaluate. One is presumably free to disagree in whole or in part with the author's thesis that much of what is now called psychology should be pruned away from from some physiological core. But the reader may find it difficult to judge to what extent, if any, the author's biases have influenced his treatment of history. Thus, several Great Men, and an even larger number of Great Theories, are treated harshly in this book. This is all right, even salutary, as long as one knows that other historians think differently.

Northwestern University

Carl P. Duncan

Thinking: From Associationism to Gestalt. By JEAN M. MANDLER and GEORGE MANDLER. New York, John Wiley and Sons, 1964. Pp. x, 300. $4.50.

In this book of selected readings, covering one line of development of the psychology of thinking from Aristotle to Duncker, the authors have interspersed comments designed to reveal the fundamental unfolding that has taken place. Stated briefly, the development of associationistic theory shows the gradual establishment of extra-mental determinants of thinking and of the correspondingly necessary distinction between the protocol statements of the introspecting subject on the one hand, and the explanatory constructs of the scientist on the other. Kurt Koffka seems to have first made the point clearly in 1912 when he argued for a separation between what he called descriptive concepts and functional concepts, the former consisting of the descriptions of conscious thought content supplied by the subject and the latter, the theoretical concepts supplied by the experimenter As an example of the confusion of these two things. Wundt conceived of attention as a conscious content discernible introspectively along with sensations and images, and yet he uses it also as a casual concept to explain the passage of impressions across the conscious threshold. In the one instance, attention appears as part of the data; in the other, it explains the data.

To develop this thesis, the authors have chosen selections beginning with Aristotle, jumping to the British associationistic school from Hobbes to Bain, crossing the North Sea to consider the Würzburg school and the discovery of imageless thought, continuing with the introduction of the explanatory concept of determining tendency by Watt, Ach, and Külpe, and finally terminating with the transition to a more purely structural explanation with Otto Selz, Koffka, Wertheimer, and Duncker. A promised second volume will contain selections from German philosophy and Husserlian phenomenology, through behaviorism, to the newly re-emerged school of French structuralists.

The defects of British associationism became evident only gradually as the doctrine reached its full elaboration at the hands of those psychologists coming at the end of the doctrinal succession—the Mills and Alexander Bain. Witness, for example, James Mill's attempt to explain the phenomenon of directedness in thinking (pp. 115-125). He discredits the theory that the mind possesses a separate power to will the course of its associations and finally describes the will as the idea of a pleasurable end—one among many ideas obeying associative bonds. Attention as a director of associations receives the same treatment, becoming (p. 123) the "occurrence of interesting sensations, or ideas." Theoretical constructs which might account for productive thinking themselves become members of the army of ideas whose directedness needs something else to explain it. In the end, Mill can give no satisfactory account of the directedness of thinking; he merely says that the discourse of some men is rambling while that of others is not, and for the latter, some "main Idea" is controlling. But how a main Idea (Will? Attention?) comes to control and in what manner it can exercise its control beyond fortuitous association is left unexplained. Perhaps nowhere else in the selection of readings is the paradox of productive thinking more strikingly revealed—how man seems to move counter to his experiences and his associations to achieve novel conclusions.

The remainder of the story is told in the history of the German psychologists. The Würzburg psychologists discovered, alarmingly enough, that introspective reports on thought revealed puzzling lacunae which did not correspond to images

or sensations. Since they were good associationists (at least in the beginning) they exercised man's eternal privilege of naming things he does not understand, dubbed the new elements *Bewusstseinlagen* (unanalyzable dispositions of consciousness), and considered them a special content of mind. The Mandlers point out with great delight how speedily these new "contents" began to appear in the introspective reports of the Würzburg protocols as *"Bsl's."* An example (p. 147). "Subject 2: '*Bsl:* Let's take the other meaning!' " (The directive factor in the thought process has now become a content)!

But the movement towards a different psychology—and a different conception of science—is already well underway. By 1904 Watt of the Würzburg school had demonstrated the vital importance of the task (*Aufgabe*) and the subject's awareness of the task upon the course of his associations; one year later Ach defines determining tendencies and demonstrates, *via* hypnotism, that they operate aside from consciousness; and in less than a decade Koffka has made the fundamental distinction between protocol statements and explanatory concepts.

Altogether the selected readings, spanning so many centuries, are remarkably coherent. One might say that the Mandlers have done an excellent job of directed historicism! But what of the future? Was such a long historical development moving towards a better psychology of thinking? The authors answer with an exuberant Yes! The field in aswarm with new ideas coming from concept formation studies, language theory, and cognitive development. Sophisticated modern theory, so far removed from mental contents that it can put them in 'black boxes,' demands testable predictions, and the handy computer stands by ready to render such demands practicable. The readings in yet a third volume are coming into existence!

Lawrence University JOHN BUCKLEW

Complex Human Behavior: A Systematic Extension of Learning Principles. By ARTHUR W. STAATS and CAROLYN K. STAATS. New York, Holt, Rinehart and Winston. Pp. xiii, 546. $7.50.

"The purpose of this book," write Staats, "is to explore various experimental and naturalistic observations of complex human behavior in terms of learning principles and thereby to offer a relatively general conception of how the physical and social environments may shape human behavior" (p. v). They also, the authors believe suffice for language development and function, for personality, human motivation, social interaction, child development, educational psychology, and behavioral problems and their treatment. An impressive level of aspiration is suggested by such topic headings as "Reinforcement and the Emergence of the Scientific Method" (p. 255). The book is intended, they say, for the introductory student or for the undergraduate major in psychology or education.

This is not the first book of its genre. Like the books of Miller and Dollard, Skinner, Osgood, Mowrer, Keller and Shoenfield—and like the articles of many others—this book is informed with a rather good idea: Let the descriptive language of classical and operant conditioning provide the theoretical language for explaining complex human behavior.

When this was a brave and new idea, some twenty-five or so years ago, it was enough for it to be reasonable and promising; but when the idea has been tried and tried, and suffered serious experimental and methodological criticism, we should

hope for a reasoned and scholarly defense, rather than a doctrinaire reaffirmation for the undergraduate reader. At the very least, we might expect the authors to temper their certitude and forebear to dismiss alternative theoretical approaches if they are not going to examine the evidence and arguments.

The authors offer their treatment of language (Chapters 4, 5, and throughout) as a principal contribution that distinguishes this book from its predecessors. Nowhere, however, do they answer, or even mention, the classical and influential indictment of their basic approach written by the linguist Chomsky in 1959. They present an essentially stochastic *S–R* model of grammatical behavior uncomplicated by any recognition of Chomsky's demonstration of the mathematical impossibility that such a model could generate the range of utterances in any normal adult's repertory. They present their own evidence for the classical conditioning of mediational meaning responses, but not the experimental re-analyses that challenge their interpretation. On Verplanck's study of human operant conditioning, they tell us that "by saying the subject 'got the idea' as to what the experimenter 'wanted' and then started doing it . . . would be to ignore the actual explanation of the subject's behavior . . . the history of reinforcement that gradually strengthens a certain type of behavior" (p. 84). They say nothing of the numerous studies supporting exactly the kind of interpretation they so flatly reject. In fact, the central contention of cognitive learning theories, that an organism responds instrumentally "because he 'knows' that if he does so he will get a (reward)," (p. 46) is listed—again without consideration of the evidence—under the heading of "Erroneous Interpretations of Learning."

These are hardly trivial lapses, since for any *S–R* analysis of reinforcement of complex human behavior, these matters are at the heart of the matter.

At the outset the authors announce that "There is no attempt to give an exhaustive account of learning principles or to consider the controversies and on-going research concerned with those that are presented" (p. v). This book, however, only pretends to be innocent of controversy. The authors close their eyes, take the student by the hand, and show him the Way.

University of Illinois

Don E. Dulaney

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review.)

Abrahams, Roger D., *Deep Down in the Jungle: Negro Narrative Folklore from the Streets of Philadelphia.* Hatboro, Pa., Folklore Associates, 1964. Pp. 287. $12.00.

Adams, Loyce, *Managerial Psychology: Human Behavior in Business and Industry.* Boston, Mass., Christopher Publ. House, 1965. Pp. xxii, 517. $9.75.

Adams, Margaret, *The Mentally Subnormal: The Social Casework Approach.* New York, Free Press, 1965. Pp. xvi, 271. $7.50.

Adler, Mortimer, J., *The Conditions of Philosophy: Its Checkered Past, Its Present Disorder, and Its Future Promise.* New York, Atheneum, 1965. Pp. xi, 302. $5.95.

Allen, Robert M., *Variables in Personality Theory and Personality Testing.* Springfield, Ill., C. C Thomas, 1965. Pp. ix, 96. $5.50.

ALLPORT, GORDON W. (ed.), *Letters from Jenny*. New York, Harcourt, Brace & World, 1965. Pp. xii, 223. Cloth, $3.50, Paper, $1.95.

AMOS, JIMMY R., BROWN, FOSTER LLOYD, and MINK, OSCAR G., *Statistical Concepts: A Basic Program*. New York, Harper & Row, 1965. Pp. ix, 125. $1.75 (paperback).

BARBE, WALTER B. (ed.), *Psychology and Education of the Gifted: Readings*. New York, Appleton-Century-Crofts, 1965. Pp. x, 534. $4.50 (paperback).

BAROJA, JULIO CARO, *The World of the Witches*. Chicago, Univ. of Chicago Press, 1965. Pp. xiv, 313. $6.50.

BERENDA, C. W., *World Visions and the Image of Man*. New York, Vantage Press, 1965. Pp. 223. $3.95.

BLOOD, ROBERT O., JR. and WOLFE, DONALD M. *Husbands and Wives: The Dynamics of Married Living*. New York, Free Press, 1965. Pp. xxi, 293. $2.45 (paperback).

BOSSELMAN, BEULAH C., ROSENTHAL, IRA M., and SCHWARZ, MARVIN, *Introduction to Developmental Psychiatry*. Springfield, Ill., C. C Thomas, 1965. Pp. v, 139. $6.50.

BRAZIER, MARY A. B. (ed.), *Brain Function. Volume II: RNA and Brain Function, Memory and Learning*. Berkeley & Los Angeles, Univ. of California Press, 1964. Pp. xvi, 360. $10.00.

BROSS, IRWIN D. J., *Design for Decision: An Introduction to Statistical Decision-Making*. New York, Free Press, 1965. Pp. xiii, 276. $2.45 (paperback).

BROWN, ROGER, Social Psychology. New York, Free Press, 1965. p. xxiv, 785. $8.95.

CANTOR, ALFRED J., *Unitrol: The Healing Magic of the Mind*. Englewood Cliffs, N.J., Prentice Hall, 1965. Pp. 192. $4.95.

CLAYTON, THOMAS E., *Teaching and Learning: A Psychological Perspective*. Englewood Cliffs, N.J., Prentice Hall, 1965. Pp. xi, 177. Cloth, $4.95; Paper, $2.25.

COHEN, JOHN (ed.), *Readings in Psychology*. New York, Hillary House Publ. Ltd., 1965. Pp. 414. $8.50.

COMMITTEE ON OCCUPATIONAL PSYCHIATRY OF THE AMERICAN PSYCHIATRIC ASSOCIATION, *The Mentally Ill Employee: His Treatment and Rehabilitation*. New York, Hoeber, 1965. Pp. 120. $2.95.

COTTY, FRED, *Basic Facts of Introductory Psychology*. New York, Macmillan, 1965. Pp. 64. $.95 (paperback).

DANIEL, ROBERT S. (ed.), *Contemporary Readings in General Psychology*. Boston, Houghton Mifflin, 1965. (First published, 1959). Pp. xiv, 417. $3.95 (paperback).

DELUCCHI, M. R., DENNIS, B. J., and ADEY, W. R., *A Stereotaxic Atlas of the Chimpanzee Brain*. Berkeley & Los Angeles, Univ. of California Press, 1965. $15.00.

DUBOIS, PHILIP H., *An Introduction to Psychological Statistics*. New York, Harper & Row, 1965. Pp. xi, 530. $7.95.

DUNNETTE, MARVIN D. and KIRCHNER, WAYNE, *Psychology Applied to Industry*. New York, Appleton-Century-Crofts, 1965. Pp. xiii, 262. $2.50 (paperback).

EBEL, ROBERT L., *Measuring Educational Achievement*. Englewood Cliffs, N.J., Prentice Hall, 1965. Pp. xii, 481. $7.95.

EBIN, DAVID (ed.), *The Drug Experience*. New York, Grove Press, 1965. Pp. 385. $.95 (paperback).

ELLIS, HENRY, *The Transfer of Learning*. New York, Macmillan, 1965. Pp. vii, 200. $1.95 (paperback).

ERICKSON, MARION J., *The Mentally Retarded Child in the Classroom*. New York, Macmillan, 1965. Pp. ix, 114. $1.25 (paperback).

EYSENCK, H. J., *Fact and Fiction in Psychology*. New York, Penguin Books, 1965. Pp. 300. $1.25 (paperback).

FEIFEL, HERMAN (ed.), *The Meaning of Death*. New York, McGraw-Hill, 1965. (First published, 1959). Pp. xvi, 351. $2.95 (paperback).

FENSHAM, PETER J. and HOOPER, DOUGLAS, *The Dynamics of a Changing Technology*. New York, Humanities Press, 1964. Pp. xvi, 248. $7.00.

Flegel, Horst. *Schizophasie in Linguistischer Deutung.* New York, Springer Verlag, 1965. Pp. 167.

Foulkes, S. H. and Anthony, E. J., *Group Psychotherapy. The Psychoanalytic Approach.* 2nd. ed. Baltimore, Penguin Books, 1965. Pp. 280. $1.25 (paperback).

Fowler, Harry. *Curiosity and Exploratory Behavior.* New York, Macmillan 1965. Pp. viii, 216. $1.95 (paperback).

Fraisse, Paul and Piaget, Jean. *Traité de Psychologie Expérimentale. IX. Psychologie Sociale.* Paris, Presses Universitaires de France, 1965. Pp. 272. F 18 (paperback).

Freeman, Paul. *An Introduction to Sigmund Freud, M.D., and Psychoanalysis.* New Jersey, Prentice Hall, 1965. Pp. unnumbered. $1.50 (paperback).

Goldstein, Henry, Krantz, David L., and Rains, Jack D. (eds.). *Controversial Issues in Learning.* New York, Appleton-Century-Crofts, 1965. Pp. ix, 476. $4.75 (paperback).

Goodenough, Erwin R. *The Psychology of Religious Experiences.* New York, Basic Books, 1965. Pp. xii, 192. $5.00.

Greene, Bernard L., *The Psychotherapies of Marital Disharmony.* New York, Free Press, 1965. Pp. xii, 191. $7.95.

Gronlund, Norman E., *Measurement and Evaluation in Teaching.* New York, Macmillan, 1965. Pp. xii, 420. $6.95.

Guilford, J. P., *Fundamental Statistics in Psychology and Education.* 4th ed. New York, McGraw-Hill, 1965. Pp. 605. $8.50.

Guion, Robert M., *Personnel Testing.* New York, McGraw-Hill, 1965. Pp. xiii, 585. $9.50.

Guirdham, Arthur, *The Nature of Healing.* New York, Hillary House Publ., 1965. Pp. 182. $5.50.

Hambling, J. and Hopkins, Philip. *Psychosomatic Disorders in Adolescents and Young Adults.* New York, Pergamon Press, 1965. Pp. ix, 246. $8.50.

Harding, M. Esther, *The "I" and the "Not-I".* New York, Bollingen Foundation 1965. Pp. x, 244. $4.50.

Harms, Ernest (ed.). *Drug Addiction in Youth.* New York, Pergamon Press, 1965. Pp. xi, 210. $9.50.

Heath, Douglas H., *Explorations of Maturity. Studies of Mature and Immature College Men.* New York, Appleton-Century-Crofts, 1965. Pp. xiv, 425. $7.50.

Henle, Paul (ed.). *Language, Thought, and Culture.* Ann Arbor Paperback 1965 (First published, 1958). Pp. vi, 273. $1.95 (paperback).

Hess, Albert G., *Chasing the Dragon. A Report on Drug Addiction in Hong Kong.* New York, Free Press, 1965. Pp. 182. $7.50.

Holt, Arthur G., *Handwriting in Psychological Interpretations.* Springfield, Ill., C. C. Thomas, 1965. Pp. xiii, 262. $10.50.

Holzkamp, Klaus. *Theorie und Experiment in der Psychologie.* Berlin, Walter DeGruyter & Co., 1964. Pp. ix, 284.

Hopkins, Philip and Wolff, Heinz (eds.). *Principles of Treatment of Psychosomatic Disorders.* New York, Pergamon Press, 1965. Pp. x, 118. [illegible]

Hovland, C. I., Lumsdaine, A. A. and Sheffield, F. D. *Experiments on Mass Communication.* New York, Wiley, 1965 (First published, 1949.) Pp. x, 345. $1.95 (paperback).

Humphreys, Christmas. *Zen. A Way of Life.* New York, Emerson Books, 1965. Pp. vii, 199. $3.95.

Imamura, Sadako. *Mother and Blind Child. The Influence of Child-Rearing Practices on the Behavior of Preschool Blind Children.* New York, American Foundation for the Blind, 1965. Pp. 78. $1.50 (paperback).

Isaacson, Robert L., Hutt, Max L. and Blum, Milton L., *Psychology: The Science of Behavior.* New York, Harper & Row, 1965. Pp. ix, [illegible]. $6.95.

Jones, Howard. *Crime in a Changing Society.* New York, Penguin Books, 1965. Pp. 174. $.95 (paperback).

KAHN, SAMUEL, *Introduction to Parapsychology.* New York, Vantage Press, 1965. Pp. 321. $5.00.
KANTOR, MILDRED B. (ed.), *Mobility and Mental Health.* Springfield, Ill., C. C Thomas, 1965. Pp. xix, 247. $10.50.
KAPLAN, LOUIS, *Foundations of Human Behavior.* New York, Harper & Row, 1965. Pp. xii, 368. $5.00.
KEMP, ROBERT, *Understanding Epilepsy.* New York, Humanities Press, 1963. Pp. vii, 92. $3.00.
KEMP, ROBERT, *Understanding Bronchitis and Asthma.* New York, Humanities Press, 1963. Pp. vii, 109. $3.50.
KESSEN, WILLIAM, *The Child.* New York, Wiley, 1965. Pp. xii, 301. Cloth, $4.95; Paper, $2.95.
KLAUSNER, SAMUEL Z. (ed.), *The Quest for Self-Control.* New York, Free Press, 1965. Pp. xiv, 400. $7.50.
KLINEBERG, OTTO and CHRISTIE, RICHARD (eds.), *Perspectives in Social Psychology.* New York, Holt, Rinehart & Winston, 1965. Pp. viii, 214. $5.25.
KLUCKHOHN, CLYDE, *Culture and Behavior.* New York, Free Press, 1965. (First published 1962). Pp. viii, 402. $2.45 (paperback).
KOTARBINSKI, T., *Praxiology.* New York, Pergamon Press, 1965. Pp. 219. $7.50.
KRASNER, LEONARD and ULLMANN, LEONARD P. (eds.), *Research in Behavior Modification: New Developments and Implications.* New York, Holt, Rinehart & Winston, 1965. Pp. viii, 403. $8.95.
LANG, ALFRED, *Über Zwei Teilsysteme der Persönlichkeit.* Bern und Stuttgart, Verlag Hans Hüber, 1965. Pp. 147.
LAWSON, REED, *Frustration: The Development of a Scientific Concept.* New York, Macmillan, 1965. Pp. x, 192. $1.95 (paperback).
LEVENSTEIN, AARON, *Use Your Head: The New Science of Personal Problem Solving.* New York, Macmillan, 1965. Pp. xiv, 270. $4.95.
LINDBLOM, CHARLES E., *The Intelligence of Democracy.* New York, Free Press, 1965. Pp. viii, 352. $7.50.
LUCE, R. DUNCAN, BUSH, ROBERT R., and GALANTER, EUGENE (eds.), *Handbook of Mathematical Psychology.* Vol. II. New York, Wiley & Sons, 1965. Pp. ix, 537. $11.95.
MARMOR, JUDD (ed.), *Sexual Inversion: The Multiple Roots of Homosexuality.* New York, Basic Books, 1965. Pp. viii, 358. $8.50.
MATHER, KENNETH, *Human Diversity: The Nature and Significance of Differences among Men.* New York, Free Press, 1965. Pp. vi, 126. $5.95.
MCNEIL, ELTON B. (ed.), *The Nature of Human Conflict.* Englewood Cliffs, N.J., Prentice Hall, 1965. Pp. xvi, 315. $6.95.
MELZER, JOHN HENRY, *Functionalism: An Outline of a Philosophy for Today.* New York, Philosophical Library, 1965. Pp. 145. $4.00.
MILLER, CARROLL H., *Guidance Services: An Introduction.* New York, Harper & Row, 1965. Pp. xi, 418. $6.50.
MONEY, JOHN (ed.), *Sex Research: New Developments.* New York, Holt, Rinehart & Winston, 1965. Pp. xii, 260. $3.50 (paperback).
MORSE, MARY, *The Unattached.* New York, Penguin Books, 1965. Pp. 226. $.95 (paperback).
MOSHER, RALPH L., CARLE, RICHARD F., and KEHAS, CHRIS D. (eds.), *Guidance: An Examination.* New York, Harcourt, Brace & World, 1965. (First published, 1962.) Pp. viii, 232. $2.50 (paperback).
MOSS, C. SCOTT, *Hypnosis in Perspective.* New York, Macmillan, 1965. Pp. x, 196. $1.95 (paperback).
MOSTOFSKY, DAVID I. (ed.), *Stimulus Generalization.* Stanford, Stanford Univ. Press, 1965. Pp. vi, 389. $11.50.
MUELLER, CONRAD G., *Sensory Psychology.* Englewood Cliffs, N.J., Prentice-Hall, 1965. Pp. viii, 120. Cloth, $3.95; Paper, $1.50.
MUSGROVE, F., *Youth and the Social Order.* Bloomington, Ind., Indiana Univ. Press, 1965. Pp. xix, 168. $5.00.

NEAL, SISTER MARIE AUGUSTA, *Values and Interests in Social Change.* Englewood Cliffs, N.J., Prentice Hall, 1965. Pp. ix, 182. $4.95.

NEWCOMB, THEODORE M., TURNER, RALPH H., and CONVERSE, PHILIP E., *Social Psychology: The Study of Human Interaction.* New York, Holt, Rinehart & Winston, 1965. Pp. xii, 591. $8.50.

OCHS, SIDNEY, *Elements of Neurophysiology.* New York, Wiley & Sons, 1965. Pp. xii, 619. $14.00.

OLDS, JAMES, OLDS, MARIANNE, BARRON, FRANK, DEMENT, WILLIAM C., EDWARDS, WARD, LINDMAN, HAROLD, and PHILLIPS, LAWRENCE D., *New Directions in Psychology II.* New York, Holt, Rinehart & Winston, 1965. Pp. x, 422. $3.50 (paperback).

PETERS, R. S. (ed.). *Brett's History of Psychology* (abridged 1 vol. ed.). Cambridge, Mass., M.I.T. Press, 1953; revised ed., 1962. Pp. 778. $3.95 (paperback).

PIÉRON, HENRI, FESSARD, ALFRED, and FRAISSE, PAUL. *L'Année Psychologique.* Paris, Presses Universitaires de France, 1964. Pp. 349.

PROKASY, WILLIAM F., *Classical Conditioning: A Symposium.* New York, Appleton-Century-Crofts, 1965. Pp. ix, 421. $7.50.

PROSHANSKY, HAROLD and SEIDENBERG, BERNARD (eds.), *Basic Studies in Social Psychology.* New York, Holt, Rinehart & Winston, 1965. Pp. xvii, 743. $8.95.

REED, ELIZABETH W. and REED, SHELDON C., *Mental Retardation: A Family Study.* Philadelphia, Saunders, 1965. Pp. v, 719. $16.50.

RICHARDSON, STEPHEN A., DOHRENWEND, BARBARA SNELL, and KLEIN, DAVID. *Interviewing: Its Forms and Functions.* New York, Basic Books, 1965. Pp. viii, 380. $7.50.

RIESMAN, DAVID (in collaboration with Nathan Glazer), *Faces in the Crowd,* abridged ed., New Haven, Yale Univ. Press, 1965. Pp. xviii, 634. $3.45 (paperback).

ROACH, JAMES H. J., SHERMAN, BARBARD R. and ROACH, PATRICIA P., *Experiments in General Psychology.* New York, Harper & Row, 1965. Pp. ix, 192. $3.75 (paperback).

ROGLER, LLOYD H. and HOLLINGSHEAD, AUGUST B., *Trapped: Families and Schizophrenia.* New York, Wiley, 1965. Pp. xiii, 436. $8.95.

ROSEN, EPHRAIM and GREGORY, IAN, *Abnormal Psychology.* Philadelphia & London, W. B. Saunders, 1956. Pp. iv, 553, $8.25.

ROSENBLITH, WALTER A. (ed.), *Sensory Communication. Contributions to the Symposium on Principles of Sensory Communication.* Cambridge, Mass., M.I.T. Press, 1961. Pp. xiv, 844. $7.50 (paperback).

ROSENBURG, MORRIS. *Society and the Adolescent Self-Image.* Princeton, N.J., Princeton Univ. Press, 1965. Pp. xi, 326. $6.50.

ROYCE, JOSIAH. *The Principles of Logic.* New York, Philosophical Library, 1961. Pp. 77. $2.75.

RUBIN, ISADORE, *Sexual Life After Sixty.* New York, Basic Books, 1965. Pp. xii, 274. $5.95.

RUITENBEEK, HENDRIK M., *The Individual and the Crowd: A Study of Identity in America.* New York, The New American Library, 1965. (First published, 1964.) Pp. 125. $.60 (paperback).

SCHILDER, PAUL, *Medical Psychology.* New York, Wiley, 1965. (First published, 1953.) Pp. 428. $2.45 (paperback).

SEVERIN, FRANK T., *Humanistic Viewpoints in Psychology: A Book of Readings.* New York, McGraw-Hill, 1965. Pp. xvii, 430.

SHERTZER, BRUCE and PETERS, HERMAN J., *Guidance: Techniques for Individual Appraisal and Development.* New York, Macmillan, 1965. Pp. x, 390. $7.30.

SHIBUTANI, TAMOTSU and KWAN, KIAN M. (Contributions by Robert H. Billigmeier, *Ethnic Stratification: A comparative approach.* New York, Macmillan, 1965. Pp. xi, 626. $7.95.

SHWAYDER, D. S., *The Stratification of Behaviour: A System of Definitions Propounded and Defended.* New York, Humanities Press, 1965. Pp. xvi, 411. $9.50.

SIDMAN, RICHARD L. and SIDMAN, MARRAY, *Neuroanatomy: A Programmed Text.* (Vol. 1.) Boston, Little, Brown & Co., 1965. Pp. x, 656. $11.50 (paperback).

SIEGEL, SIDNEY (in collaboration with Alberta Engvall Siegel, and Julia McMichael Andrews), *Choice, Strategy, and Utility.* New York, McGraw-Hill, 1965. Pp. xii, 189. $5.95.

SIMON, HERBERT A., *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization.* New York, Free Press, 1965. (First published, 1945.) Pp. xlviii, 259. $2.45 (paperback).

SIMPSON, GEORGE EATON and YINGER, J. MILTON, *Racial and Cultural Minorities.* 3rd. ed. New York, Harper & Row, 1965. Pp. xiii, 582. $8.75.

SKINNER, B. F., *Science and Human Behavior.* New York, Free Press, 1965. (First published, 1953.) Pp. x, 461. $2.95 (paperback).

SLUCKIN, W., *Imprinting and Early Learning.* Chicago, Aldine Publ. Co., 1965. Pp. x, 145. $5.00.

SPEERS, REX W. and LANSING, CORNELIUS, *Group Therapy in Childhood Psychosis.* Chapel Hill, N.C., Univ. of North Carolina Press, 1965. Pp. xviii, 186. $6.00.

SPOCK, BENJAMIN and LERRIGO, MARION O., *Caring for Your Disabled Child.* New York, Macmillan, 1965. Pp. viii, 373. $4.95.

STEIN, MAURICE, VIDICH, ARTHUR J., and WHITE, DAVID MANNING (eds.), *Identity and Anxiety: Survival of the Person in Mass Society.* New York, Free Press, 1965. Pp. 658.

STEINER, IVAN D. and FISHBEIN, MARTIN (eds.), *Current Studies in Social Psychology.* New York, Holt, Rinehart & Winston, 1965. Pp. xii, 532. $7.95.

STEPHENS, JOHN M., *The Psychology of Classroom Learning.* New York, Holt, Rinehart & Winston, 1965. Pp. xiv, 511. $7.50.

STOODLEY, BARTLETT H. (ed.), *Society and Self: A Reader in Social Psychology.* New York, Free Press, 1965. (First published, 1962). Pp. 713. $4.50 (paperback).

STOUFFER, S. A., SUCHMAN, E. A., DEVINNEY, L. C., STAR, S. A., and WILLIAMS, R. M., JR., *The American Soldier.* Vol. 1, 2. New York, Wiley & Sons, 1965 (first published, 1949). Pp xii, 599. $2.95 (paperback), Vol. 2. Pp. 675, $2.95 (paperback).

TATE, MERLE W., *Statistics in Education and Psychology. A First Course.* New York, Macmillan, 1965. Pp. ix, 355. $5.95.

TIFFIN, JOSEPH and MCCORMICK, ERNEST J., *Industrial Psychology.* 5th Ed. Englewood Cliffs, N.J., Prentice Hall, 1965. Pp. xiv, 682. $8.95.

TINGSTEN, HERBERT, *The Problem of Democracy.* Totowa, N.J., Bedminster Press, 1965. Pp. 210. $5.00.

VYGOTSKY LEV SEMENOVICH (Trans. by Eugenia Hanfmann, and Gertrude Vakar, *Thought and Language.* Cambridge, Mass., M.I.T. Press, 1962. Pp. xiii, 168. $2.45 (paperback).

WAPNER, SEYMOUR and WERNER, HEINZ (eds.), *The Body Percept.* New York, Random House, 1965. Pp. x, 112. $1.65 (paperback).

WEISMAN, AVERY D., *The Existential Core of Psychoanalysis: Reality Sense and Responsibility.* Boston, Mass., Little, Brown & Co., 1965. Pp. xii, 268. $7.50.

WHITTAKER, JAMES O., *Introduction to Psychology.* Philadelphia & London, W. B Saunders, 1965. Pp. xxiii, 631. $7.50.

WHITTAKER, JAMES O., *Instructor's Manual For Use With Introduction to Psychology.* Philadelphia & London, W. B Saunders, 1965. Pp. v, 62.

WHITTAKER, JAMES O., *Student's Workbook to Accompany An Introduction to Psychology.* Philadelphia & London, W. B. Saunders, 1965. Pp. viii, 224. $3.25 (Paperback).

WISDOM, J. O. and WOLFF, HEINZ (eds.), *The Role of Psychosomatic Disorder in Adult Life.* New York, Pergamon Press, 1965. Pp. xiv, 229. $8.50.

YOUNG, ROBERT K. and VELDMAN, DONALD J., *Introductory Statistics for the Behavioral Sciences.* New York, Holt, Rinehart & Winston, 1965. Pp. x, 435. $7.95.

THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

VOL. LXXVIII DECEMBER, 1965 No. 4

A COHERENCE-CRITERION IN PERCEPTION

By ALBERT S. RODWAN, Emory University

In perceptual research, it is generally assumed that *O*s act to satisfy some subjective criterion. Traditionally, this criterion has been assumed to be accuracy; more correctly, it has been assumed that *O*s maximize the correlation between stimuli and responses. It is suggested here that a more critical and adequate criterion is coherence, the way in which an *O* makes sense out of the stimulus-variation. This can be, at times, accuracy but need not be. Furthermore, in many situations, an *O* must satisfy a dual criterion: his own and *E*'s. In those cases in which no criterion of accuracy exists, or is known, knowledge of the *O*'s criterion becomes crucial to understand his perception.

There is some evidence to show that criteria other than accuracy can exist. Rodwan and Hake have shown that when accuracy is not defined, *O*s do consistently use certain responses when presented with certain stimuli.[1] This implies that *O*s can hold, in a consistent manner, standards as criteria which are known only to themselves. Their study dealt with the judged intelligence of schematic faces and he showed that he could account for approximately 95% of *O*'s judgments when assuming a coherence-criterion.

The mode of analysis was a linear discriminative function (*LDF*) which required *O* to partition a set of stimulus-objects into two classes which are then used to infer the existence of two experiential classes. As a result,

* Received for publication November 9, 1962. This work was supported in part by Grant G 16352 from the National Science Foundation to H. W. Hake. This work was done at the University of Illinois.

[1] A. S. Rodwan and H. W. Hake, The linear discriminative function as a model for perception, this JOURNAL, 77, 1964, 380-392.

the *LDF* can be considered as a psychophysical method; and it has valuable properties: there is no response-variance; it can take into account *n* dimensions; and it produces a set of weights which are indices of the importance of each of the physical dimensions.

On the basis of the partitioning of the stimuli, a linear function is derived which maximizes the distance between the two experiental classes on a subjective axis. This subjective axis is called a 'decision-axis'. The coefficients in the linear function are the indices of the importance of each of the subjective dimensions in the total decision.

There are certain assumptions involved in deriving the *LDF* and they are given by Rodwan and Hake.[2] It may be advantageous to recapitualte them: The first assumption is that when a stimulus has *n* physical dimensions and *O* takes into account all *n*, then that stimulus has *n* psychological dimensions and can be considered to be a point in a subjective *n*-space. The second assumption is that *O* uses a criterion of maximization in perception. That is to say, the *LDF* assumes (this is just a mathematical assumption) that *O* uses that subjective decision-axis which maximize the ratio of the variance between the two experiental classes to the variance within the two classes. The last assumption is that the variance within the two classes is the same. This last assumption is a convenience only in that the discriminative function is linear if the two variances are equal; if they are not equal, then a quadratic discriminant will satisfy the rest of the assumptions.

Another aspect of the *LDF* is that it provides a theoretical basis for a study of perception. The one additional step needed to make the *LDF* a theory of perception is to coördinate the mathematical assumptions with critical psychological assumptions. The one mathematical assumption which is critical is that of maximization.

There are three different ways of stating the assumption of maximization. One is to maximize the ratio of the variance between the two classes of experience to the variance within the classes. The second is to maximize the distance, on some subjective dimension, between the means of the two classes of experience. The third is to maximize the correlation between the elements of the two classes of experience on some subjective dimension, and a linear combination of the physical dimensions of the stimuli.

These three statements of the maximization are equivalent and, therefore, that subjective dimension which satisfies one, will satisfy all three. It has been shown that the *LDF* represents that subjective dimension which maximizes the ratio of between-classes variance to within-classes variance; furthermore, the *LDF* is that subjective dimension which produces a maximal correlation. The weights of the *LDF* are proportional to the regression-weights in the multiple point biserial regression-equation.

Therefore, that the *LDF* may be a theory of perception, one of the three equivalent mathematical statements must make psychological sense. The correlational statements fits most naturally in that it says that *O* must maximize the correlation between real stimulus-variation and his experience, r_c. That is to say, if the *LDF* is an adequate theory of perception, then the *O*s must act as if they always made the maximal sense, coherence, out of the stimuli. The maximization is independent of any external criterion such as an accuracy.

[2] Rodwan and Hake, *op. cit.*, 384.

Hake presented an argument to the effect that whatever the perceptual system is, it is not a fidelity-system.[3] He argued that the perceptual system does not maximize accuracy, but rather that it maximizes coherence. Brunswik argued the same thing when he said that the perceiver maximizes his perceptual achievement.[4]

The usual definition of coherence is in terms of processes or functions. Suppose two processes or functions are denoted by f and g; the statement that "two processes are coherent" is to be understood to mean that a correlation exists between them. Thus if f is a function defined on the stimulus-class and g is a function defined on the set of responses, and if a correlation exists between stimuli and responses, *i.e.* between f and g, then we would say that the two classes were coherent. Thus, we see that accuracy is a special case of coherence, the correlation between stimuli and responses. On the other hand, coherence would exist if f denoted the set of experiences of the observer and g denoted the class of stimuli and if f and g were correlated. This type of coherence can exist even though there is no accuracy, *i.e.* even though the correlation between the stimuli and responses is zero. Thus, coherence is a more general term than accuracy and, as used here, refers to the way in which an *O* makes sense out of the variation in the physical stimuli. When an *O* interprets the stimulus-variation as variation of a particular sort, he, in effect, establishes a correlation between the stimulus-variation and an internal standard. To this internal standard of variation we give the name "standard of coherence." For example, the *O* may be presented with a set of plane figures varying in w/h, the width-height ratio, with a constant area; he can then 'experience' this set of stimuli as having a constant w/h, but varying in area. In this case, his standard of coherence is areal variation and the variation experienced makes sense even though a correlation computed between the stimuli and responses is zero.

Internal standards have been discussed in the psychological literature. Helson has suggested that stimulus-variation is interpreted as variation above and below a null point.[5] He states that a standard of coherence is provided by this null point and the subjective dimension on which it (the null point) exists. The existence of stimulus-anchors have been also suggested.[6] Stimulus-anchors are points on the stimulus-dimension which have fixed corresponding points on the *O*'s subjective dimension; *O*'s experiences are interpreted with respect to these latter points. In this case, the anchor-points and the subjective dimension on which they exist supply the standard of coherence.

Under the assumption that *O* acts as if he used the *LDF* in making his judgments, that is, as if he defined coherence as the correlation between the stimulus-variation and the variation on the *LDF*, r_C, and as if the criterion of maximization were a standard against which to compare his experiences, we can make certain predictions about the outcome of an experiment in which the criterion accuracy is partially destroyed.

[3] H. W. Hake, Contributions of psychology to the study of pattern vision, *WADC* tech. *Rep.* TR 57-621, 1957.

[4] Egon Brunswik, *Perception and the Representative Design of Psychological Experiments*, 1956.

[5] Harry Helson, Adaptation theory, in Sigmund Koch (ed.), *Psychology: A Study of a Science*, 1958.

[6] C. W. Eriksen, and H. W. Hake, Absolute judgments as a function of stimulus range and number of stimulus and response categories, *J. exp. Psychol.*, 49, 1955, 323-332; and F. C. Volkmann, and Trygg Engen, Three types of anchoring effects in the absolute judgment of hue, *J. exp. Psychol.*, 61, 1961, 7-17.

The experiment completed required *O*s to judge a set of stimuli which consisted of plane figures (rectangles). Two judgments were required: a judgment of form ("Is the presented stimulus-figure a square or a rectangle?") and a judgment of size ("Is the presented stimulus-figure large or small?"). The range of variation of the height and width among the rectangles was made so small that little accuracy could be achieved with respect to the real characteristics of the form. It was predicted that even though *O*s could not satisfy a criterian of accuracy very well, they would act to satisfy a criterion of coherence defined by the correlation between the stimuli and *O*'s experiences, r_C. That is, the *O*s were expected to satisfy this purely subjective criterion in a consistent way. If they did so, then for each *O* an *LDF* should predict responses accurately, even though there was no accuracy *i.e.* even though r_{SR} is zero. In addition, two distinct *LDF*'s should apply: one for the form-judgments and one for the size-judgments. Judgments were made for several durations of viewing to test whether, for a particular kind of judgment, the *LDF* showed stability within the *O*s. Also, if the *LDF* is to have and generality, it should reflect the use of the same criterion of coherence over durations of viewing which occur in every day perception.

METHOD

Apparatus and stimuli. The apparatus was a rack-and-pinion device which generated squares and rectangles whose sides varied from zero to 100 mm. This device was set in one arm of a modified Gerbrands tachistoscope with a light-source behind it so that an *O* would see a rectangle of light surrounded by a dark field. The viewing distance was 8 ft. A fixation-point was timed to go on when the back light went off. The fixation-light was centered between the left and right edges of the stimulus-figure and was 1 in. above the tallest figure. Viewing was monocular, with the left eye. The light-source for the stimulus was two 4-w. fluorescent bulbs and for the fixation-light it was a single 4-w. bulb. The fixation-bulb was so shielded that it was just discernable upon looking into the apparatus.

The stimulus-set consisted of 100 figures composed of 10 different heights and 10 different widths. Both height and width varied from 79.0 mm. to 80.8 mm. in increments of 0.2 mm. Thus, the set consisted of two classes of figures, squares and rectangles, each of which varied in area. The stimulus-set was an orthogonal set (one in which the dimensions along which variation occurred were all orthogonal to one another).

Precedures and obseriers. The judgments required of the *O* were four, two single-binary and two double-binary decisions. The two single-binary decisions were: (1) to decide whether the presented figure was a square or a rectangle, or, (2) to decide whether the presented figure was large or small. The two double-binary decisions were: (1) to decide on the form first and then on the size, or, (2) to make the size-decision first followed by the form-decision. The second decision in the double-binary decisions was always made on the same stimulus-presentation.

All of these judgments were required for each value of the duration of viewing parameter *t*. The values of *t* were 0.01, 0.05, 0.10, 0.20, 0 30, 0.50, 0.75, 1.00, 1.50, and 2.50 sec.

Three undergraduates served as *O*s, two were women and one a man. None had ever participated in a perceptual experiment prior to this one. Each *O* underwent four preliminary trials (combination of judgments and value of *t*) that he might become familiar with the experiment. The first preliminary condition was a form-judgment with the duration of viewing at 1.00 sec.; the second preliminary was a size-judgment at the same duration of viewing; the third was a double-binary judgment, with form first and size second; the fourth condition was the reverse of the third. Both of these were at durations of 1.00 sec. In the experiment proper, the 40 conditions were randomized, as were the 100 stimuli in each condition. At the beginning of each condition, *O* was told which decision was required and the duration of viewing for that condition. The first 20 presentations were practice trials, after which the experimental stimuli were presented. *O* knew that the first 20 trials were practice in that he was not asked to record them but was asked to record the subsequent 100 responses. *O* was not told whether the presented figure was a square or rectangle, or large or small.

The intertrial-interval (the time from the onset of the stimulus to the onset of the subsequent one) was constant for the entire experiment, at 3.5 sec. One condition was completed in 20 min. The total number of experimental hours was 22, which included the preliminary conditions. Thus, two conditions were run during each hour.

Each *O* was told that he was participating in a psychophysical experiment. *E* explained that 'psychophysics' was a method of determining what it was that *O* saw when he was presented with stimuli. *O*s were told that since *E* was interested in what they, the *O*s, saw, there was no correct or incorrect answer. Two of the three *O*s immediately voiced suspicions. They reported that this was the standard set of instructions given students who were participating in 'personality-type' experiments. These two *O*s were offered the task of aiding *E* in calibrating the apparatus and randomizing the condition and trials. They were also asked to speak to their friends, individuals who had taken a class in experimental psychology, about psychophysics. Only after each *O* was satisfied that he was, in fact, participating in a psychophysical experiment, was he allowed to begin. The third *O* was immediately satisfied in that he had just completed a class in experimental psychology. All three *O*s reported, after the first preliminary condition, that there were too many tall figures; when questioned, they said that they had seen too many, not that *E* had changed the stimulus-set on them. When reminded that *E* had not changed the stimulus-set, they 'knew' that there were not too many tall figures, and the experiment continued. The *O*s were convinced of *E*'s lack of guile.

Results

The first result to be reported concerns the accuracy with which the *O*s identified the stimuli. Accuracy of judgment was expected to be low because of the very small differences in height and width which distinguished the stimuli.

Accuracy was low for both form- and size-judgments, and was measured by means of the now familiar *d'* measure of signal detection-theory.[7] This has the critical advantage of providing an estimate of the acuity of discrimination which is uncontaminated by the rate at which *O* chooses to detect one of the stimulus-classes. For both form- and size-judgments, *d'* measures were computed within *O*s for each duration of viewing. The measures were unrelated to duration of viewing and varied approximately

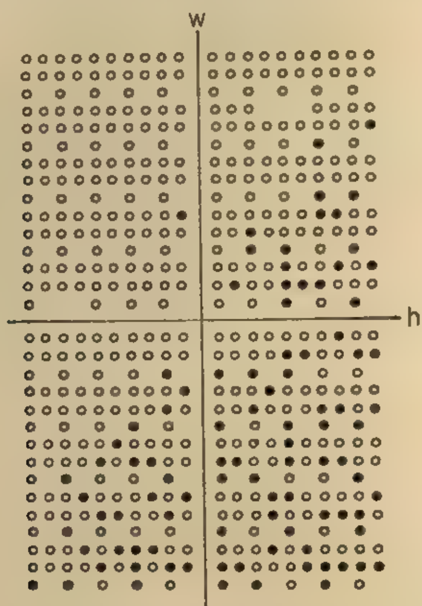

Fig. 1. A Sample of Five Form Judgments for the Following Values of *t*: 0.01, 0.05, 0.75, 01.50, and 2.50 Sec. (The open circles denote a "rectangle" response and the filled, a "square" response. The data are from *FA*.)

from −0.50 to 0.50, the average value being about zero. This result indicates the existence of little or no accuracy of judgments when accuracy is defined in terms of the physical characteristics of the stimuli. In the case of form-judgments, the response "square" could be applied accurately to only 10 stimuli; in the case of the size-judgments, one half of the stimuli were large. The *O*s were told which half were the large and which were

[7] J. A. Swets, W. P. Tanner, Jr., and T. G. Birdsall, Decision processes in perception, *Psychol. Rev.*, 68, 1961, 301-304.

small. For the form-judgments, actual squares occurred only rarely; the ratio of squares to rectangles was 1:10. The ratio of the response 'rectangle' to the response 'square' was about 8:2. The *O*s used the response "square" too often. The ratio of the responses "large" to "small" was about 1:1.

The way in which the *O*s used their responses to partition the stimuli into two classes for each kind of judgment is shown in Fig. 1 for the form-judgments and in Fig. 2 for the size-judgments. In each figure, the

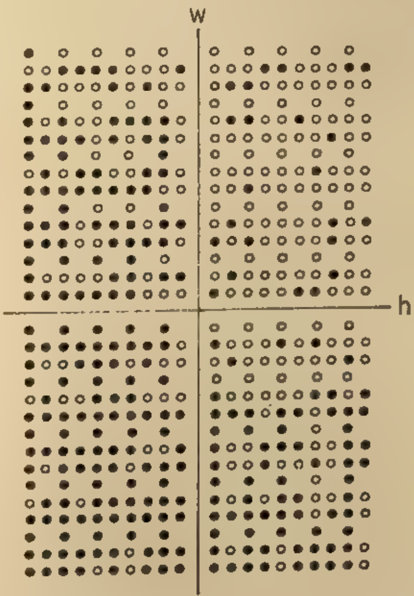

FIG. 2. A SAMPLE OF FIVE SIZE JUDGMENTS FOR THE FOLLOWING VALUES OF *t*: 0.01, 0.05, 0.75, 1.50, AND 2.50 SEC. (The open circles denote "large" and filled denote "small" judgments. These data are for *FA*.)

vertical dimension represents the *width* of the stimulus-figures and the horizontal dimension represents their height. Each stimulus-figure is represented by a cell in this matrix of stimulus-figures. Within each cell are five circles representing the responses of *FA* for each of the five durations of viewing: 0.01, 0.05, 0.75, 1.50 and 2.50 sec. For each trial in which this *O* called a stimulus-figure a "rectangle," an open circle occurs in that cell. Closed circles denote the occurrence of "square" responses to appropriate stimuli. The same representation of the size-judgments of *FA*

occurs in Fig. 2, in which open circles denote "large" responses and closed circles "small" responses.

In Fig. 1, the actual partitioning of the stimuli into squares and rectangles departs radically from the partitioning demanded by criterion of accuracy. To be accurate, *O* should have applied the response "square" to stimuli on or along the positive diagonal of the stimulus-matrix shown in the figure, and all other stimuli should have been called "rectangle." Instead, *O* established one class of figures, those above the positive diagonal, which he uniformly called rectangles; stimuli below the positive diagonal were indiscriminately called squares and rectangles; which response *O* used was unpredictable from the duration of viewing or physical dimension of the stimulus. In the case of size-judgments, Fig. 2, the behavior of *O* was somewhat more straightforward. Stimuli above the negative diagonal tended to be called "large" and those below the negative diagonal "small." Although not apparent in Fig. 2, *FA* showed some of the asymmetry of responses indicated in Fig. 1. That is, the conditional probability of calling a large stimulus "large" was higher than the conditional probability of calling a small stimulus "small."

In summarizing results, which are typical of those of the other *O*s, it can be said that the *O*s did not use, or did not satisfy, an accuracy-criterion when they were judging the stimuli. They did partition the stimuli into two classes; but this partitioning except perhaps for the size-judgments, shows no straightforward, clear relationship to the actual or physical characteristics of the set of stimuli. This is especially true of the form-judgments in which one homogeneous class of stimuli, called rectangles, and another, called either square or rectangle, was established.

Two points are raised by the results. One is the possibility that *O*s, especially in the case of the form-judgments, were acting to satisfy a subjective criterion, *i.e.* making private sense out of the stimuli. The nature of this criterion cannot however, be determined from the analysis thus far. The implication of the results concern the *LDF:* The nature of the partitioning evident in Fig. 1 and 2 would indicate that the *LDF* probably cannot serve, in any simple fashion, as an adequate model for all judgments made by the *O*s.

The mechanisms for deriving the *LDF* are given by Rodwan but a brief recapitulation here is in order.[8] A set of weights, a vector, for a linear function is required which maximizes r_c. If the stimuli are orthogonal, the vector, V, is given by:

$$V = kM_r, \qquad [1]$$

where V is the required vector, k is a constant of proportionality, and M_r is the the vector of means for the rectangles, in deviational scores. It is important to

[8] Rodwan, *op. cit.*, 384f.

note that M_r is the vector of means for the class of subjective rectangles, *i.e.* for the class of figures which *O* assigns to the rectangular class. They are not the physically defined rectangles. If only the relative weights are required, we can solve for k as follows:

$$V'V = k^2 M'_r M_r, \text{ and}$$
$$k = 1/(M_r' M_r)^{\frac{1}{2}}. \qquad [2]$$

Tables I, II, and III give the relative (*i.e.* normalized) weights for the *LDF*'s for each *O*, *FA*, *NC* and *JK* respectively.

There are three aspects of Tables I, II, and III which are of interest. First, there seems to be no effect due to the duration of viewing. Secondly, there is a difference between the decisional axes of form and size but it is one of sign only and not of magnitude. The sign of the height coefficient is positive for the size-axes but negative for the form-axes. Thirdly, within

TABLE I

NORMALIZED LDF'S FOR *O* FA FOR ALL CONDITIONS
(The decimal point has been omitted from each number.)

	Form alone		Form first		Form second		Size alone		Size first		Size second	
	w	*h*	*w*	*h*	*w*	*h*	*w*	*h*	*w*	*h*	*w*	*h*
0.01	943	−333	916	−402	707	−707	879	478	946	324	892	450
0.05	982	−188	908	−419	877	−481	894	−447	840	543	964	−265
0.10	919	−394	674	−739	999	−001	983	−182	999	001	974	225
0.20	768	−640	546	−838	965	−263	836	548	916	401	839	544
0.30	447	−894	867	−498	962	−273	928	−371	816	578	687	727
0.50	399	−917	797	−604	951	−310	941	339	995	099	889	459
0.75	912	−410	965	−263	701	−713	993	−118	774	633	976	219
1.00	887	−461	775	−632	798	−605	802	598	949	316	999	016
1.50	915	−494	983	−181	881	−473	987	161	998	040	995	096
2.50	809	−588	514	857	994	−447	874	486	642	767	735	678

either the judgments of form or size there is no effect due to order of responding; that is, there seems to be no effect which shows a dependence on which judgment came first nor any difference between the single and double judgments. Each table can be considered as being homogeneous within types of judgments. Fig. 6 is a frequency distribution of the angle that the decision-axis (*LDF*) makes with the height-axis. Each point is for a different value of *t*. It can be seen that there is very little overlap between size-and form-axes. In general, the size-axes have a positive weight for height, while form-axes have a negative one. There are certain exceptions, but they are not crucial. The positive sign for the *h* coefficient means that the decision-axis was above the *h*-axis, while a negative sign means that it was below the *h*-axis.

It appears, so far, that if *O* responds to anything, *i.e.* to some criterion, that criterion remains relatively constant within the two classes of judgments. Fig. 3, 4, and 5 show that the *O*s did, as a matter of fact, respond meaningfully.

Figs. 3, 4, and 5, for *Os FA, NC,* and *JK* respectively, give the conditional probability of the *LDF* assigning a stimulus to a category of response that the *O* assigned it to that category. Figs. 3A, 4A, and 5A give the conditional probability for the form judgments, while Figs. 3B, 4B, and 5B give the conditional probability for the size judgments. The dotted line is for the single judgment, the dashed line is for the form (size) first, and the solid line is for the form (size) second. The upper curves in each figure are for the rectangle judgments (Figs. 3A, 4A, and 5A) or for the large judgments (Figs. 3B, 4B and 5B). These three figures show that the *LDF*'s are extremely good for rectangle and large judgments, but cannot distinguish the other two cases. A second point of interest is that the decision-axes are approximately constant for all values of *t*.

TABLE II

NORMALIZED LDF's FOR *O* NC FOR ALL CONDITIONS
(The decimal point has been omitted from each number.)

t	Form alone		Form first		Form second		Size alone		Size first		Size second	
	w	*h*	*w*	*h*	*w*	*h*	*w*	*h*	*w*	*h*	*w*	*h*
0.01	174	−985	976	−212	−604	−797	406	914	707	707	772	635
0.05	981	−196	871	−492	879	−478	304	953	493	763	900	435
0.10	678	−735	346	−938	938	−347	665	747	757	654	707	−707
0.20	581	814	995	099	999	−001	958	287	907	420	882	471
0.30	287	−758	629	−777	678	−733	996	−087	971	−123	826	563
0.50	183	−232	043	−999	973	−229	901	433	694	720	833	553
0.75	−316	−949	518	−855	−530	−848	667	745	677	736	179	984
1.00	034	−998	124	−992	374	−927	916	401	609	793	406	914
1.50	707	−707	949	316	409	−899	665	756	814	581	335	942
2.50	090	−996	998	053	707	−707	799	601	747	664	903	429

TABLE III

NORMALIZED LDF's FOR *O* KJ FOR ALL CONDITIONS
(The decimal point has been omitted from each number)

t	Form alone		Form first		Form second		Size alone		Size first		Size second	
	w	*h*	*w*	*h*	*w*	*h*	*w*	*h*	*w*	*h*	*w*	*h*
0.01	974	−224	026	−999	151	−989	862	507	324	946	832	555
0.05	602	−799	713	−646	811	−585	724	690	489	873	625	781
0.10	894	−447	519	−855	870	−494	946	329	−063	998	801	598
0.20	814	−581	678	−735	548	−836	208	778	408	913	224	974
0.30	884	−467	752	−659	351	−936	552	834	725	689	316	949
0.50	935	−354	910	−414	894	−447	838	546	669	744	597	802
0.75	860	−510	772	−635	670	−742	396	918	366	931	482	876
1.00	794	−607	830	−557	707	−707	769	640	643	766	503	864
1.50	796	−606	671	742	788	−616	853	757	496	868	656	754
2.50	885	−466	688	−726	530	−848	597	802	880	475	521	854

As a general statement about these data, it can be said that the *LDF* is showing *O* to be using a consistent criterion and that this enables him to separate the stimuli into two classes only one of which is homogeneous. This applies to both judgments, but more clearly to the form judgment.

DISCUSSION

These data demonstrate that there is little or no criterion of accuracy operating, but that the *Os* are satisfying some subjective criterion and are doing so in a consistent manner; the agreement between the *LDF* and the *O*'s responses for the "rectangle" and "large" classes shows that they are, in part at least, satisfying *rc*.

The nature of this result, *i.e.* the agreement between the *O*'s responses and the

prediction of the *LDF*'s for only one of two classes, raises some interesting questions about the behavior of *O*'s perceptual system. In the case of form judgments, there are two possibilities. The partitioning of the stimuli of Fig. 1 means that the *O*s clearly reorganized all the stimuli above the positive diagonal as rectangles, but that the stimuli below the positive diagonal constituted a region of uncertainty, making a diffuse boundary between what *O* would call a square and what he

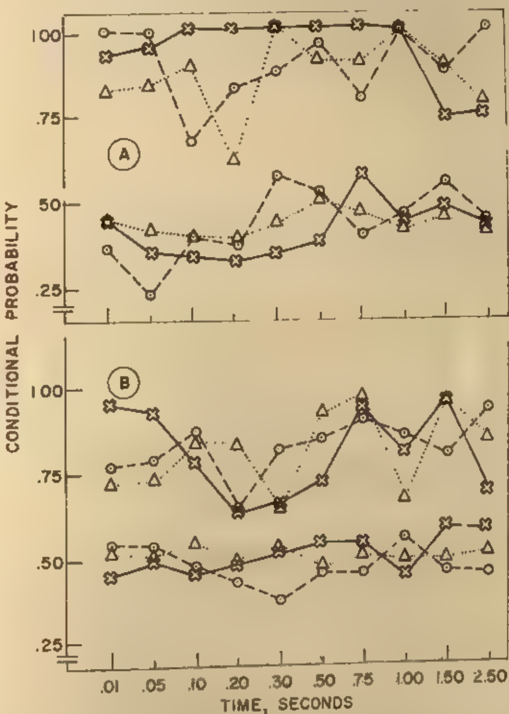

FIG. 3. THE CONDITIONAL PROBABILITY THAT THE *LDF* ASSIGNED A STIMULUS TO THE SAME RESPONSE CLASS AS DID *FA*.
(See text for further discussion.)

would call a rectangle. The boundary between what is called square and what is called rectangle does not correspond to physical squareness. Rather, it corresponds to a large set of stimulus-figures which are taller than they are wide. The implication is that these stimulus-figures are shortened, in experience, in the vertical dimension. Tall stimuli are likely to be seen as square; this constitutes a 'negative' vertical-horizontal illusion. As a result, the stimuli experienced by the *O*s did not consist of a symmetric set. The ratio of stimuli experienced tall to those experienced as short was less than unity. This raises the question: Does this illusory effect exist in the apparatus? One possible answer was given by the *O*s themselves. Their com-

plaint, reported above, was that they saw too many tall stimuli, not too many squat stimuli. Unfortunately, a question about the nature of the distortion in the apparatus is difficult to answer. As a preliminary to the data reported here, extensive work was done in the laboratory to learn the form of the stimulus which an *O* would call "square." Square, as an attribut of plane figures, was discovered to be an elusive concept; the stimulus which an *O* would call square varied with the psychophysical method, as well as within and between experimental sessions. The variations measured, in

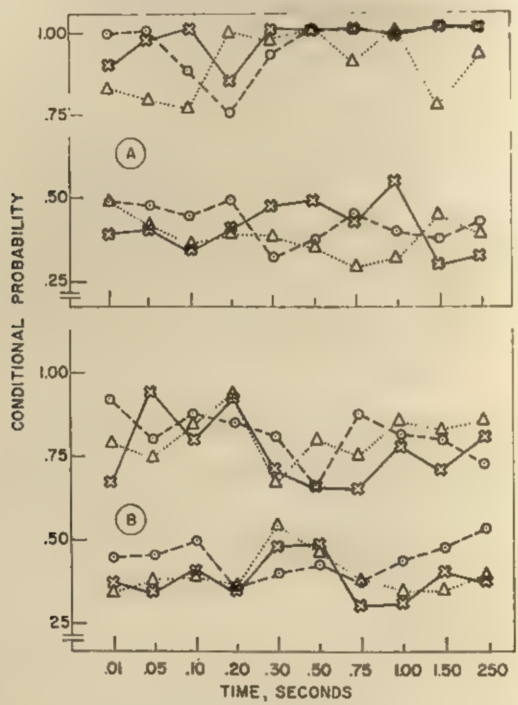

Fig. 4. The Conditional Probability That the *LDF* Assigned a Stimulus to the Same Response Class as Did *NC*.

most cases, were at least double the range of variation represented by the stimuli used in this experiment.

It should be noted that the present experiment represents a prototype of all such experimentation; that is, this experiment is a paradigm for determining the subjective values required for an *O* to be satisfied that a stimulus-object is "square." A set of stimuli differing from squares by being too tall or too short is presented to an *O* who is asked to respond "square" to that stimulus which appears to be a square. As a study designed to learn about the nature of the illusory character of this situation, the present one appears to have a single unusual aspect; namely, the range of variation allowed in the set of stimuli was extremely small. The im-

plication appears clear: if a greater range of variation in the stimuli had been allowed, *O*s could have established a better partitioning of the stimuli. Square figures and tall figures would have been called "rectangle" and a set of figures somewhat taller than wide would have been called "square."

This suggestion, as a real possibility, depends on two assumptions: (a) the existence of a negative vertical-horizontal illusion; and (b) the fact that a greater

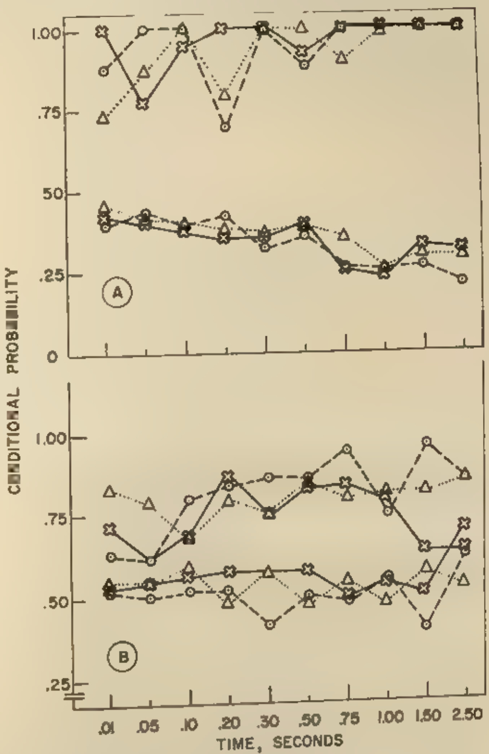

FIG. 5. THE CONDITIONAL PROBABILITY THAT THE *LDF* ASSIGNED A STIMULUS TO THE SAME RESPONSE CLASS AS DID *JK*.

range of stimulus-variation would include more of the much taller figures which *O* would have to call "rectangle."

A counter-possibility is that the use of a greater range of stimulus-variation would not lead to such an unequivocal result. The reasonableness of the counter-suggestion is predicated on the fact that the range of stimulus-variation used was rather large. It can be shown that *O*s can discriminate a rectangle (the tall narrow one used in this experiment) from a square of the same width and the same square from a rectangle which is lower than the square (the short one in this experiment) much better than chance. The *d*s from such experiments are in the

4.0 to 5.0 range. In fact, it is difficult to build an apparatus which can produce differences smaller than *O*s can reliably discriminate.

A second aspect of the counter-possibility is that the use of intermediate stimuli is expected to 'jam' this acuity.[9] The critical question being asked is: how do *O*s react when a discriminable difference, *i.e.* that between the tallest and shortest rectangles, is filled with intermediate stimuli? The *LDF* suggests both a method by which the action of *O*s can be understood in this situation, and an hypothesis about what *O* would do: They will partition the stimuli to achieve the largest possible difference between two experienced classes.

This attempt was only partially successful. In the case of form-judgments, especially, the *LDF* is capable of predicting responses reliably to only one class of stimuli. The *LDF* analysis, however, could tell us, for that class, how the *O*s weighted the stimulus-dimensions in their definition of that class; and the *LDF*'s

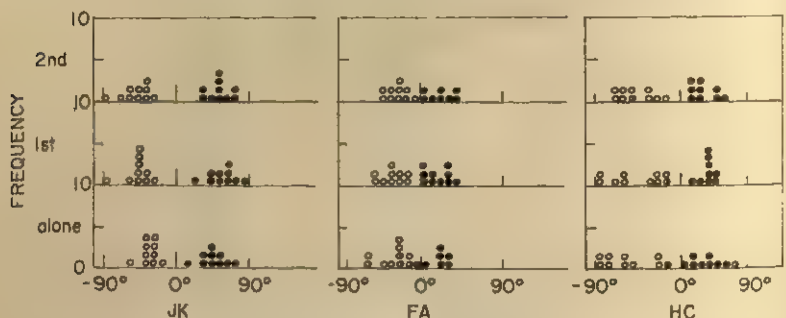

Fig. 6. Frequency Distributions of the Angle Whose Cosine (Arc Cosine) is the Coefficient of the Dimension of Height

for the form-judgments were definitely different from those for the size-judgments. There is another suggestion as to why the *LDF* was successful for only one class of stimuli. This suggestion involves the cut-offs used by the *O*s.[10] Cut-offs are the subjective cutting lines used by *O*s when they partition the stimuli into two classes of experience. Experience on one side of the cut-off would be called by the name of one category of response and those on the other side would be called by the name of the other category. The assumptions of the *LDF* methodology is that this cut-off is stable for an *O* after many trials have been run. To say the same thing differently, the *LDF* assumes that experience becomes stablized. It is possible, however, that this is not true. Conceivably, two states can exist for one set of stimuli: (a) *O* sees stimuli in the upper left of Fig. 1 and his cut-off is somewhere near the positive diagonal. This leads to his consistently calling those stimuli "rectangles." This is a stable state over trials. The second state (b) is when *O* sees stimuli in the lower right of Fig. 1. This is an unstable state and the instability arises because *O* (as reported earlier) experiences too many tall figures. Since he

[9] D. E. Broadbent, *Perception and Communication*, 1958, 27.
[10] Swets, Tanner, and Birdsall, *op. cit.*, 306.

knows that the rectangular stimuli are symmetrical (half are tall and thin, while half are short and squat), he tries to compensate for whatever process produces his asymmetric experiences. On some trials, he must accept a tall-appearing stimulus as being the result of a process of distortion. To do this he must move his cutting line, *i.e.* his cut-off, toward the lower right in Fig. 1. By so doing, he accepts a tall-appearing stimulus-figure as a "square" and calls it a square. On other trials, he accepts tall-appearing stimuli as rectangles. In those cases, his cut-offs remain somewhere near the positive diagonal of Fig. 1. In this unstable state (the state where *O* sees tall figures), his cut-off becomes a variable and has at least two values over trials. The result is the partitioning of Fig. 1 in which there are one homogeneous class of stimuli and one heterogeneous class.

In conclusion, then, two suggestions have been made about the data: (1) Because of the limited range of variation of the stimuli and because of a distortion was operating to make square figures look squat, *O*s could establish only one clear-cut class of squat figures bounded by a wide zone of uncertainty. This zone of uncertainty corresponds to tall figures which appear to be squat. (2) *O*s could deal quite easily with squat figures which they could call "rectangles." Because, however, of a process of distortion, which tended to make figures appear taller than they were wide, *O*s experienced more tall figures than they thought they should have and tried to compensate for this by using a shifting cut-off (at least two cut-off placements) in that case where they did not trust their experience.

The existence of these counter-arguments suggests, of course, further research. This should involve the use of stimuli covering a wider range of variation. Predictions from the counter-arguments appear particularly clear for an experiment in which the range of stimulus-variation is increased only by adding taller figures, that is, increasing the size of the matrix of Fig. 1 below the positive diagonal.

The first suggestion predicts that *O*s should establish two classes of experienced rectangles (tall and squat) separated by a rather broad region of uncertainty corresponding to all the figures which are only a little taller than they are wide. These would be called squares. The second suggestion leads to the prediction that this would lead to an even greater asymmetry of experience with many more taller figures than squat figures being experienced. In that event, the *O*s could be expected to try to compensate to an even greater degree by the use of an unstable cut-off. Stability of the cut-off, in fact, should be achieved only if the set of stimuli presented were changed by the addition of more squat figures; that is, by extending the matrix in Fig. 1 to the left of the positive diagonal, to produce a symmetric set of experienced figures. In which case, squat and tall figures would be experienced about equally often.

Summary

An experiment was designed to determine whether the linear discriminant function (LDF) could serve as a model for perceptual psychophysics as well as a theory of recognition. The results indicated: (a) that an accuracy-criterion was not met; (b) that the *LDF*'s for form and for size were clearly different; and (c) that the *LDF* is an adequate model for predicting one class of responses (in the form-judgment it was rectangle and in the size-judgment it was large). Two counter-suggestions were considered in attempting to account for these results and the nature of the future research was described.

SINGLE-TRIAL LEARNING: A STOCHASTIC MODEL FOR THE RECALL OF INDIVIDUAL WORDS

By RAND B. EVANS and KARL M. DALLENBACH, The University of Texas

Since the postulation of the Traditional Laws of Association, authors have considered 'repetition' or some form of it, such as 'frequency' or 'exercise,' as the primary basis of learning. Gestalt psychology, however, holds that this is in error; it maintains that learning is 'insightful,' that it occurs suddenly at a single trial—and numerous studies have been reported in support of this point of view.[1] Since these results do not fall within the scope of the traditional pattern, learning theorists have either ignored them completely or have regarded them as Gestalt phenomena and thus outside their field of concern.

Rock's demonstration of single-trial learning, which falls within the mainstream of contemporary objectivism may not, however, be so cavalierly ignored.[2] Since his 'break through,' a large number of experiments, reporting results pro and con, have appeared. In most of these investigations, the method of paired-associates—the method employed by Rock—was used. As results are a function of the method and as the method of paired-associates may be conducive to single-trial learning, it is, as Clark, Lansford, and Dallenbach point out, highly desirable that other methods of learning be used in investigating this problem.[3] Murdock and Babick, accordingly, used the method of free recall in their investigation of the problem of single-trial learning.[4]

The use of the method of free recall is important. There is a great difference between the recall of the association between a pair of words and the recall of an individual word. The method of paired-associates is a rather complicated process. *O* must first recognize the initial word or syllable in the matched pair and then must recall the word or syllable with which it is paired in previous exposures. It would seem that the recall of individual words or syllables is a much more basic

* Received for publication June 14, 1964. The authors are indebted to Professors John Theios and W. K. Estes for their assistance on the stochastical aspects of the study.

[1] Wolfgang Kohler, *Gestalt Psychology*, 1929, 371 f., 376 ff.; cf. G. L. Kreezer and K. M. Dallenbach, Learning the relation of opposition, this JOURNAL, 41, 1929, 432-441; Mary Schooley and G. W. Hartman, The role of insight in the learning of logical relations, *ibid.*, 49, 1937, 237-292; J. I. Lacey and K. M. Dallenbach, Acquisition by children of the cause-effect relationship, *ibid.*, 52, 1939, 103-110.

[2] Irvin Rock, The role of repetition in associative learning, this JOURNAL, 70, 1957, 186-193.

[3] L. L. Clark, T. G. Lansford, and K. M. Dallenbach, Repetition and associative learning, this JOURNAL, 73, 1960, 22-40, esp. 38.

[4] B. B. Murdock, Jr., and A. J. Babick, The effect of repetition on the retention of individual words, this JOURNAL, 74, 1961, 596-601.

and simple kind of learning than that of paired-associates. Since a scientific psychology operates on the procedure of beginning with the simplest and most unitary principles and then building on them the more elaborate and involved, it would seem that the recall of individual words is a particularly important method to use in investigating the tenability of all-or-none learning.

Murdock and Babick did, indeed, find that something like all-or-none learning existed in the recall of individual words; at least, their data did not differ significantly from a single parameter model of all-or-none learning which predicts a geometric distribution of trials before the first recall.[5]

The procedure used by Murdock and Babick was imaginative and effective, but it has serious shortcomings in its control of several factors: (1) One is the reading of the stimulus-words to *O* by *E*. This method of presentation is subject to several errors: (a) *E* may unconsciously emphasize some words over others; (b) the verbal presentation of the stimulus-words and the intervals between successive presentations are not as uniform or as well controlled as visual presentation. This discrepancy is particularly great when multi-syllabic words are used together with monosyllabic words. (2) Another shortcoming was in the use of the same single-exposure words more than once. Under the hypothesis advanced in their study, this procedure was not permissable. The effect may have been negligible, but when carrying out an exploratory study in such an uncertain field, one never knows what effect such a contradiction might have.

The purpose of the present study is threefold: (1) to repeat the Murdock and Babick study, using stricter controls; (2) to expand the basic procedure used to gain information about certain aspects of the learning process which Murdock and Babick did not consider; and (3) to construct a mathematical model that can predict the results of these two experiments and that may be generalized to other studies within the field.

Materials. The stimulus-words were presented visually on white cold-press surface, illustration-board cards, 11.5 x 6 cm. in size. The cards were 2 mm. thick to prevent crushing or bending when presented in the apparatus. The words were printed on the cards in India ink with a lettering set, and the letters were 1.2 x 1.2 cm. in size. The cards were then sprayed with plastic to preclude any spurious cues' being given by soiling or otherwise marking the cards. The stimulus-words were selected at random from a list of four-letter, one-syllable meaningful words taken from the book of common English words by Thorndike and Lorge.[6] All homonyms were removed.

Apparatus. The apparatus consisted of a timer-controlled mechanical card changer. The front of the apparatus consisted of a 55 x 50 cm. screen, in which a 7.5 x 4 cm. window was cut and through which the exposed words were viewed. The apparatus was painted flat black. The screen concealed *E* from *O*.

The words were exposed for 0.6 sec., with intervals of 0.4 sec. between successive words, making a complete cycle between successive words of 1 sec. These

[5] *Ibid.*, 596-601.

[6] E. L. Thorndike and Irving Lorge, *The Teachers Word Book of 30,000 Words*, 1944.

intervals were chosen to prevent rehearsing that might take place during the long exposure-times used by Rock and many of his successors. More importantly, however, the short exposure-times were chosen to prevent the invention of bizarre associations or other mnemonic devices as aids to recall. As Clark Lansford, and Dallenbach have shown, mnemonic devices are utilized by the *O*s when exposure-times and interexposure intervals exceed 1 sec.[7] The importance of controlling such associations is even greater when it is realized what their effect is on the computation of an average among *O*s. These associations are extremely variable and idiosyncratic. Thus, one *O* might make one type of association to remember one group or pair of words in the serial order, another might create another type of association for the same words or for different words, and another *O* might make no association at all. The large numbers of possible associations available in a serial list make the averaging of the results of the *O*s under such qualitatively widely divergent conditions difficult and really of little meaning. Clark, Lansford, and Dallenbach showed that the reduction of exposure-time to 1 sec. effectively reduced the use of mnemonic aids. The use of 0.6 sec. in this experiment seems to have virtually eliminated such devices, since no *O* reported being able to use them. Realistically, of course, it must be recognized that the past experience of *O* will make some words easier to recall than others. Also, it should be realized that there are probably some associations being created between the various words in the serial list. The main purpose of the short exposure-time was to eliminate conscious systems or mnemonic aids.

The experiments were conducted in a quiet room. The sounds that were present (air conditioning, generator hum) and the sounds of the apparatus, were constant and masked the outside noises.

EXPERIMENT I

Observers. Ten *O*s, both men and women, were obtained from Freshman courses in psychology. The data from one *O* were discarded without computation because of his report that he used a conscious system, a violation of instructions. Thus only 9 *O*s were considered in the results.

Instructions. The following instructions were given the *O*s:

You will be shown a series of words, after which, at the signal, 'recall,' write as many of the words as you remember on the sheet of paper provided. You will have 1 min. in which to do this. At the end of that time, hand me the paper and a series will again be exposed. Attend to every word as it is shown to you. That is, do not concentrate on some words and slight others. Have you any questions?

Procedure: The *O*s were tested individually. Each *O* was seated about 80 cm. from the apparatus. *O* was shown 30 series of cards with 15 words per series. Three classifications of words were prepared and used: non-experimental words, single-exposure words (corresponding to Murdock and Babick's non-critical words), and multiple-exposure words (corresponding to Murdock and Babick's critical words). The non-experimental words were those appearing in serial positions 1–5 and 11–15. The single-exposure words occupied four of the five central serial positions (Positions 6 through 10). The multiple-exposure words occupied one position, randomly placed in Serial Positions 6 through 10.

The non-experimental words and the single-exposure words were shown only

[7] Clark, Lansford, and Dallenbach, *op. cit.*, 32-37.

once to each *O*, and were changed on every series. The multiple-exposure word was repeated series after series, until recalled. Then a new multiple-exposure word was inserted in the next list and the procedure was repeated until the 30 series had elapsed. The multiple-exposure words were maintained in the same serial position until recalled, the next multiple-exposure word being placed in another serial position. As far as possible, all of the serial positions, Positions 6 through 10, were used equally often. The order of the lists containing the non-experimental and single-exposure words was different for each *O*. The order of the multiple-exposure words remained the same for all *O*s, but since different *O*s took different numbers of exposure to recall the various multiple-exposure words, the effect was that of seldom having the same multiple-exposure word appear in the same single-exposure word list for any two *O*s.

The choice of the five central positions in the serial list for experimental purposes was made because it is here that the serial position curve is the flattest. The non-experimental words were used to take care of the effects of recency and primacy in the recall of serially-presented words.

The method of free recall was used. *O*s were given 1 min. in which to write as many words as could be recalled. This time seemed sufficient, since *O*s seemed to have always finished by the end of that time. There was a lapse of 15 sec. between the time *O* handed the recalled list to *E* and the beginning of the next list. Since *O* realized that the list, or most of it, would not be exposed more than once, the problem of rehearsing between trials is believed to be negligible. After the fifteenth series, *O* was given a 2-min. rest-period. There were varying numbers of multiple-exposure words shown to different *O*s, since the replacement of a multiple-exposure word was dependent upon its recall. Eight different multiple-exposure words were shown during the experiment, with an average of 2.74 exposures required for recall with a variance of 0.59.

Results. The first matter of interest was to determine if the single-exposure words were different in their difficulty of recall from the multiple-exposure words. The comparison was made by checking the probability of recalling the single-exposure words with that of recalling the multiple-exposure words exposed for the first time. The mean probabilities were 0.215 and 0.250, respectively. The difference was far from significant ($t = 0.264$, $df = 8$). It may be considered, then, that the multiple-exposure words and the single-exposure words are of nearly equal difficulty.

Following Murdock and Babick's analysis, the next consideration was to see what effect repetition has on the probability of recalling individual words. Murdock and Babick hypothesized that if repetition has no significant effect on the probability of recall, then the distribution of the probabilities of the recall of multiple-exposure words should match a geometric distribution, PQ^{x-1}, where P is the probability of recalling the word on an exposure. Q is the probability of not recalling the word on that exposure, and x is the number of times the word is exposed. If the

hypothesis that repetition facilitates the recall of words is correct, the observed and theoretical distributions should not fit, but the distribution of the probability of recalling multiple-exposure words should diverge farther and farther from the theoretical distribution as the number of exposures increases. This geometric distribution is a single-parameter model for all-or-none learning. Like Murdock and Babick, the *P*-value chosen for the creation of the distribution was that of the probability of recall of the single-exposure words. The results may be seen in Fig. 1.

It will be noted that there is a peak after the second exposure on the observed distribution. This is a result that the geometric distribution does

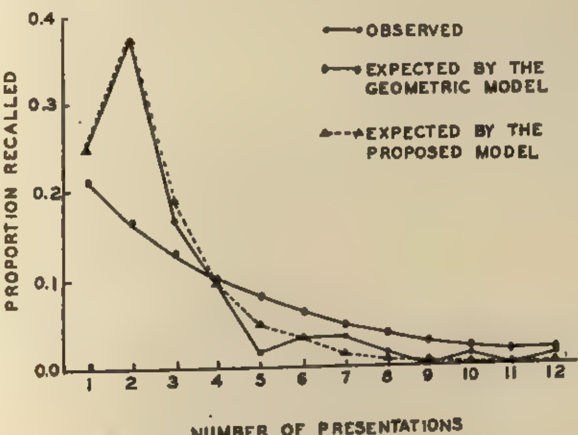

FIG. 1. PROPORTIONS OF EXPERIMENTAL WORDS RECALLED
Experiment I

not, and cannot, predict. The goodness of fit between the hypothesized and observed distribution was assessed with the Kolmogorov-Smirnov test.[8] A maximal cumulative deviation of 0.268 was obtained, which is significant at the 0.05 level of significance with 72 cases.

The problem of averaging, which Murdock and Babick encountered, was not a problem here, since there were no significant differences among the multiple-exposure words in difficulty at the 0.05 level of significance. An analysis of variance resulted in an F of 1.71, $df = 8{,}56$. Thus, the value of the probability of recall of single-exposure words may be considered representative of the probability-values for the various multiple-exposure words. The results seem to deny the Murdock and Babick model.

[8] Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences*, 1956, pp. 47-52.

The observed distribution significantly differs from a geometric distribution and there are definite patterns of deviation that the geometric model does not predict. Dissatisfaction with this failure of the geometric model to predict experimental results led to the creation of a mathematical model to better predict these results. The model will be discussed later.

EXPERIMENT II

There are some aspects of the learning process in the recall of individual words that are desirable to investigate that cannot be studied by using the Murdock-Babick method as it stands. It is also desirable to study the tenability of all-or-non learning in other situations.

Observers. Nine *Os* were obtained from Freshman courses in psychology. Both men and women were used.

Instructions. The instructions were the same as in the first experiment.

Procedure. New words were drawn from the remaining pool of four-letter, one-syllable, non-homonym words. The procedure was similar to that of the first experiment. As before, three types of words—non-experimental, single-exposure, and multiple-exposure—were used. The non-experimental words again took serial positions, Positions 1–5 and 11–15. In this experiment, however, the number and position of the single-exposure words and the multiple-exposure words were reversed, giving four of the central five positions to the multiple-exposure words in each list and one of the five positions to a single-exposure word. Single-exposure words were changed on each series. The non-experimental words and the multiple-exposure words were exposed series after series, until all four of the multiple-exposure words had been recalled. Then a new non-experimental word and multiple-exposure word-list were exposed. The position of the single-exposure words remained constant in a particular multiple-exposure list until the multiple-exposure list was changed. So far as was possible, the serial positions, Positions 6–10 were used equally often. The non-experimental words and multiple-exposure words in each list were shown in the same order to each *O*, but the lists themselves were arranged in different orders for all *Os*. The single-exposure words were shown in the same order for each *O*.

O was shown 30 series of 15 words per series, as in the first experiment. There was a rest-period of 2 min. after the 15th series. The method of free recall was used. Forty multiple-exposure words were shown to each *O* in all and these were repeated until recalled. An average of 2.60 exposures was required for the first recall. The variance among the 40 words was 0.56.

Results. The assumption made in Experiment I and in the Murdock and Babick study was that the central serial positions of the word-lists were the locations of the flattest portions of the serial-position curve from the data at hand and that there was no significant difference in the probability of recall of the central words. With the procedure followed in this experiment, it was possible to determine empirically the effect of

serial position on the recall of the central words. The curve of serial position for these data was in the classic form, with increasing ease of recall the farther the word from the center. There was a marked serial position-effect from Positions 1–15. The central positions (Positions 6–10) were, as was hypothesized, the flattest portion of the curve. There were no significant differences in difficulty of recall of the multiple-exposure words in the various central positions. An analysis of variance produced an F of 0.519, $df = 5, 32$.

As in the previous experiment, *E* wished to test if the single-exposure words differed in difficulty of recall from the multiple-exposure words. To accomplish this, the probability of recalling single-exposure words was compared to the probability of recalling multiple-exposure words after their first exposure. The probability of recalling the single-exposure words was somewhat higher than that of recalling the multiple-exposure words on their first exposure, but the difference was not significant at the 0.05 level of significance ($t = 1.45$, $df = 8$).

Again following the analysis of Murdock and Babick, the influence of the repetition of words on the probability of recall was tested. The distribution of the probabilities of recalling multiple-exposure words after different numbers of exposures was compared to a geometric distribution constructed from PQ^{x-1}, where P was the probability of recalling a single-exposure word. The observed distribution again differed significantly from the all-or-none geometric model, producing a Kolmogorov-Smirnov $D_{max.} = 0.19$. With 244 cases, this is significant at the 0.05 level of significance. Fig. 2 shows the observed distribution and the geometric model. It will be noted that again there is a marked peak after the second exposure. In a cumulative distribution, this shows up as a systematic cumulative deviation of the observed distribution from the geometric distribution over the middle numbers of exposures.

The problem of non-homogeneity among words and *O*s that troubled Murdock and Babick was, again, not encountered here. A comparison of the word requiring the fewest repetitions for recall and the word requiring the most repetitions for recall for the most *O*s produced a t of 2.50, df 5. This was not significant at the 0.05 level.

In this experiment, the single-exposure words were changed every series. In this way, the single-exposure words were presented in multiple-exposure lists that had been presented from 0 to 9 times. The question arises as to whether this adding of a new word to a list exposed several times has any effect of drawing attention to the new word.

Fig. 3 shows the results of this analysis. It will be noted that there is

a general increase in the probability of recalling a single-exposure word as the number of times the list on which it is exposed as been exposed. On the second exposure of the repeated lists, however, the probability

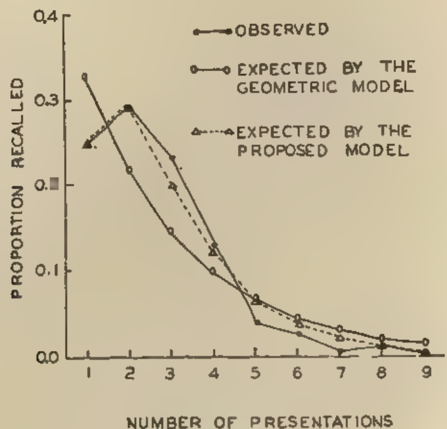

FIG. 2. PROPORTIONS OF EXPERIMENTAL WORDS RECALLED
Experiment II

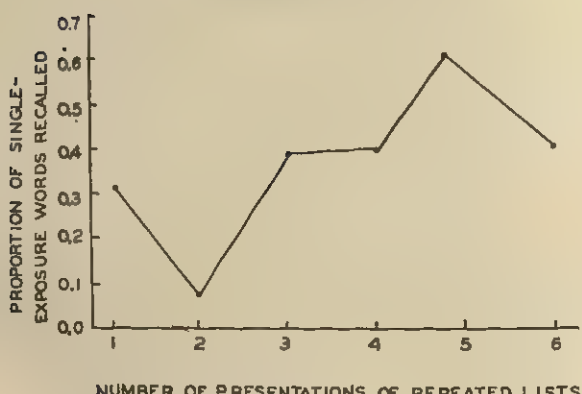

FIG. 3. PROPORTIONS OF WORDS RECALLED

of recall of single-exposure words drops to a very low value. It would seem that attention is actually being drawn away from new words placed in the lists after the lists had already been exposed once, but that this effect lessens more and more in subsequent exposures. This finding takes on added significance in the context of the model to be described.

Model and Discussion

The model presented here and shown in Fig. 4 is similar to one used by Waugh and Smith and works in the context of the general model for learning of Bower and Theios.[9] The model may be represented as a three-state, two-parameter Markov chain. In this model, words are either stored or not stored on a given exposure, with a probability *c* and *1-c*, respectively. The words that are stored may be stored in one of two different places on an individual exposure. The words may go directly into recall storage, in which case the word will be recalled for the first time, with a probability *d*, which will be assumed to be equal to *c*. The word may go into a recognition-storage, with a probability *1-c*. The words in the

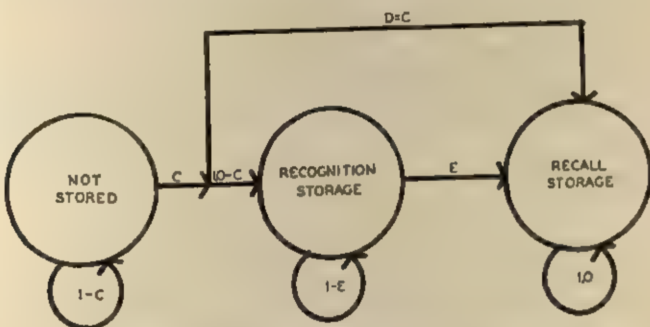

Fig. 4. The Proposed Three-State, Two-Parameter Markov Model for the Recall of Individual Words

recognition-storage may transfer into recall storage on the following exposure with a probability *E*. Since there can be no words in recognition-storage carried over into recall until after their second exposure, the probability of recall after the first exposure is c.[2] If *F* represents the number of exposures before first recall, the value of *c* may be estimated by:

$$\hat{c} = \sqrt{Pr(F = 1)} \qquad [1]$$

where $Pr(F = 1)$ is taken as the observed proportion of items recalled after the first exposure.

The value of the transfer from recognition-storage into recall-storage, *E*, may be estimated by the formula:

[9] Nancy Waugh and J. E. K. Smith, A stochastic model for free recall, *Psychometrika*, 27, 1962, 141-152; G. H. Bower and John Theios, A learning model for discrete performance levels, in R. C. Atkinson (ed.), *Studies in Mathematical Psychology*, 1963, 1-32.

$$\hat{E} = [Pr(F = 2) - c^2(1 - c)]/[c(1 - c)] \qquad [2]$$

This model and the general model for learning devised by Bower and Theios may be considered equivalent by taking $P = 0$ and $d = c$ in the latter model. Since this is the case, Equation [32] given by Theios may be adapted to the present model by taking $P = 0$ and $d = c$.[10] Thus, the probability of recalling words for the first time on a given exposure K may be given by:

$$Pr(F = K) = c^2(1 - c)^{K-1} + \frac{(1 - c)cE}{(1 - E) - (1 - c)}[(1 - E)^{K-1} - (1 - c)^{K-1}] \qquad [3]$$

This probability-distribution gives all the data of a free verbal recall experiment of the Murdock and Babick type, and thus it is a sufficient statistic with which to test any model for first recall.

This model is merely a corollary of the general Bower and Theios model. Likewise, the model devised by Waugh and Smith is a special case of the model proposed here. In terms of this model, Waugh and Smith assumed $c = E$. It should be mentioned here that one estimation-equation

$$\hat{E} = (1 - c)/[F - 1/c] \qquad [4]$$

which follows from the Theios equation was not used for the estimation of E, since the mean number of repetitions before the first recall which is F in the formula turned out not to be a good estimator of E.[11] Equation [2] was found to be a far superior estimator for the data examined.

It is important to remember that this model predicts the probability of a first recall of words. No assumptions or predictions are made concerning subsequent recall after the first recall.

By the application of the model to the first experiment of this study, by Formulas [1] and [2], a $c = 0.5$ and $E = 1.00$ were obtained. Fig. 1 demonstrates how closely the model fits the observed data. The Kolmogorov-Smirnov test of goodness of fit applied to these data results in a maximal cumulative deviation ($D_{max.}$) between the two distributions of 0.05. The single-parameter geometric distribution gave a Kolmogorov-Smirnov $D_{max.}$ of 0.268 for these same data.

When the model was applied to the second experiment, where c was estimated as 0.5 and E as 0.66, another close fit was obtained. The Kolomogorov-Smirnov $D_{max.}$ was 0.045. Fig. 2 demonstrates how close is the

[10] Bower and Theios, *op. cit.*, 1-32.

[11] John Theios, The mathematical structure of reversal learning in a shock-escape T-maze: Overtraining and successive reversals, *J. math. Psychol.* 2, 1965, 26-52.

fit. The geometric model produced a Kolmogorov-Smirnov $D_{max.}$ of 0.19 for these same data.

To demonstrate further the generality of the model, the results of Waugh and Smith were analyzed by this method, with $c = 0.5$ and $E = 0.46$. Fig. 4 demonstrates that the fit here is also very close, producing a Kolmogorov-Smirnov $D_{max.}$ of 0.027, as compared to a fit of 0.037 using the formula prepared by Waugh and Smith.

It is striking that three separate experiments, each using slightly different methods and one from a different laboratory should produce a c of 0.5. There is no particular reason that this should be, as the model is not bound to this one value of c.

The data from an earlier pilot study were analyzed to demonstrate this.

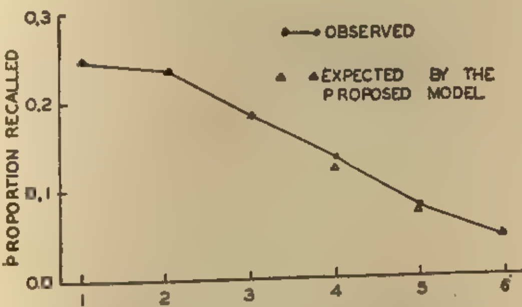

FIG. 5. WAUGH AND SMITH'S DATA COMPARED WITH THE PREDICTIONS MADE FROM THE PROPOSED MODEL $c=0.50$ and $E=0.46$

There were eight *O*s in this experiment, and the procedure was the same as that of Experiment II, except that the multiple-exposure words were changed whenever the single-exposure word that appeared in a series was recalled. Symbol c was estimated as 0.63 and E as 0.827. Again, the Kolomogorov-Smirnov test was used, and a $D_{max.}$ of 0.0166 was obtained —a close match.

Although no formal introspective technique was carried out in this study, several of the *O*s reported that they noticed words disappearing from the list, but no *O* reported noticing a word's being added to the list. This suggests either a combing through the lists for words that may be recognized or that there is an increased sensitivity to some words in the list rather than to others. Reference to Fig. 3 will show that words appearing for the first time in lists that were being exposed on the second exposure had a very low probability of recall, 0.078. The probability of

recalling a new word added to a list that was being exposed for more than two times becomes greater and greater. It is interesting to note that the greatest amount of transfer from recognition- to recall-storage occurs during the second exposure of the words and then decreases on subsequent exposures. These results, plus the reports of the *Os* of disappearing words, suggests that words in recognition-storage, though not recallable, are more recognizable than new words in the list. Attention may be drawn to these words, with recognition in the transference of the words into recall-storage. It may be that the average effectiveness of the recognition of these recognition-storage words is a determinate of the value of *E*.

It should be noted that the geometric model referred to requires that the parameter *P* remain constant over trials. The analysis shown in Fig. 3 indicates rather directly the fact that the parameter *P* does not remain constant over trials. It is possible that a modified geometric model that took this variation into account could provide better prediction than one put forth by Murdock and Babick.

Although this study refutes the model of the learning process for free recall presented by Murdock and Babick, it does not refute the basis of their assumption that learning is an all-or-none process; indeed, our results confirm theirs. The findings of this study further confirm Rock's studies in that repetition does not seem to have any effect on the acquisition of words for the first recall, beyond the allowance of more chances for a word to be acquired in an all-or-none fashion. The present model also seems to indicate that the storage-processes after word-acquisition and up to the first recall, function on a fixed percentage-basis, which is basically an all-or-none process. This model makes no assumptions or predictions about recall of particular words after their first recall.

PITCH-DISCRIMINATION AT HIGH FREQUENCIES BY AIR- AND BONE-CONDUCTION

By JOHN F. CORSO, State University of New York, Cortland, and MURRAY LEVINE, Fort Sam Houston

In the psychophysical study of human hearing, the acoustic stimuli may be presented by either of two primary routes: (1) the usual one by way of the external auditory meatus, designated as air-conduction; and (2) the more uncommon one by way of the bones of the skull, designated as bone-conduction.

Several earlier studies have pointed to similarities between these two modes of transmission. Von Békésy showed that when a 4,000 ~ tone was presented to the forehead, suitable changes in the intensity and in the phase of a binaural air-conducted tone of the same frequency produced a cancellation of the tone.[1] Cancellation-effects have also been shown by Lowy for tones from 250–3000 ~ and by Wever and Lawrence for tones from 100–1500 ~ .[2] Wever and Bray, measuring cochlear potentials, found that the two forms of stimulation exhibit the same functional relationship to the intensity of the stimulus.[3] It appears, therefore, that the two modes of acoustic transmission involve the same sensory cells with the same patterns of displacement in the cochlea, regardless of whether the stimulation is applied by air or by bone conduction.

More recent data, however, suggest that unsuspected differences between the systems of air- and bone-conduction may perhaps exist. In terms of the frequency-range of auditory sensitivity, most investigators agree that the human ear can respond to air-conducted tones from approximately 15–20,000 ~ .[4] Wever extends the upper limit to approximately 24,000 ~ for young people with no otological impairment and Corso and Oda obtained data to 25,000 ~ .[5] Pumphrey reported that his subjects responded only to tones below 16,000 ~ .[6]

---

*Received for publication August 12, 1963. This study was performed at The Pennsylvania State University supported by grant B-2006(R1) to the senior author from the National Institute of Neurological Diseases and Blindness. The authors are indebted to Dr. Fujio Oda and Mr. John L. Searle for technical assistance.

[1] Georg von Békésy, Zur Theorie des Horens bei der Schallaufnahme durch Knochenleitung, *Ann. d. Physik*, 13, 1932, 111-136.

[2] Karl Lowy, Cancellation of the electrical cochlear response with air- and bone-conducted sound, *J. acous. Soc. Amer.*, 13, 1942, 156-158; E. G. Wever and Merle Lawrence, *Physiological Acoustics*, 1954, 234.

[3] E. G. Wever and C. W. Bray, The nature of bone conduction as shown in the electrical response of the cochlea, *Ann. of Otol. Rhinol. Laryngol.*, 45, 1936, 822-830.

[4] R. S. Woodworth and Harold Schlosberg, *Experimental Psychology*, 1956, 331.

[5] E. G. Wever, *Theory of Hearing*, 1949, 330; J. F. Corso and Fujio Oda, Preliminary report on upper frequency audibility functions, *Res. Bull. No. 17, Dept. of Psychol., Pennsylvania State Univ.*, 1961.

[6] R. J. Pumphrey, Upper limit of frequency for human hearing, *Nature*, 166, 1950, 571.

For bone-conduction, the literature is mainly concerned with the clinical frequency-range of 250–4000 ~ . Data beyond these limits are given by Watson who indicated a possible range from 20–20,000 ~ and by Frazier and Watson who obtained thresholds from 80–12,000 ~ .[7] The findings in a series of recent studies however, extend the upper limit of audibility by bone-conduction well beyond that conventionally given for air-conduction. Pumphrey obtained auditory sensations by bone-conduction up to 100,000 ~; Deatherage, Jeffress and Blodgett reported thresholds for an underwater tone of 50,000 ~ ; Corso and Oda have reported a threshold-curve for bone-conduction from 5,000–100,000 ~ ; and Haeff and Knox have obtained data from 20,000–108,000 ~ .[8] Thus, the available data on the range of sensitivity indicate a possible difference in the two modes of acoustic transmission and suggest that the cessation of sensation by air-conduction at the upper end of the audibility-range is not the result of the simultaneous failure of the conductive and neural processes.

A comparison of the air- and bone-conduction with respect to the upper limit of the perception of pitch suggests a second difference between the two modes of transmission. For air-conduction, Pumphrey has found that discrimination begins to fail above 12,000 ~ ; Wever and Lawrence state that audible frequencies up to 24,000 ~ produce a single pitch; Ward reports that, for trained musicians, pitch disappears at about 5500 ~ .[9] For bone-conduction, Pumphrey concluded that pitch ceased to be discriminated at approximately 16,000 ~ ; Deatherage, Jeffress, and Blodgett believe that pitch was not perceived in the ultrasonic region and that the pitch of an ultrasonic tone is comparable to the highest tone an individual can hear; Corso and Levine have confirmed this observation in reporting that tones of 57,000, 64,000 and 94,000 ~ are matched in pitch to a tone of approximately 17,000 ~ .[10] Thus, it seems that the upper limit of pitch-perception for pure tones is somewhat higher for bone- than for air-conduction, but additional data are needed on this matter.

The present study was designed to investigate further the differences between air- and bone-conduction. Specifically, the primary purpose was to establish a psychophysical function for pitch-discrimination for the two modes of transmission from 2000 ~ to approximately 14,000 ~. In addition, the study was intended to provide data on pitch-discrimination for a 57,000 ~ tone presented by means of bone-conduction.

---

[7] N. A. Watson, Limits of audition for bone conduction, *J. acous. Soc. Amer.*, 9, 1938, 294-300; T. V. Frazier and N. A. Watson, Bone conduction threshold measurements, *ibid.*, 20, 1948, 220 (Abstract).

[8] Pumphrey, *op. cit.*, 571; B. H. Deatherage, L. A Jeffress, and H. C. Blodgett, A note on the audibility of intense ultrasonic sound, *J. acous. Soc. Amer.*, 16, 1954, 582; Corso and Oda, Bone-conduction thresholds for sonic and ultrasonic frequencies, *ibid.*, 34, 1962, 746 (abstract); A. V. Haeff and Cameron Knox, Perception of ultrasound, *Science*, 139, 1963, 590-582.

[9] Pumphrey, *op. cit.*, 571; Wever and Lawrence, *op. cit.*, 23; W. D. Ward, Subjective musical pitch, *J. acous. Soc. Amer.*, 26, 1954, 369-380.

[10] Pumphrey, *op. cit.*, 571; Deatherage, Jeffress, and Blodgett, *op. cit.*, 582; John Corso and Murray Levine, The pitch of ultrasonic frequencies heard by bone conduction, *Proceedings of the Pa. Acad. Sci.*, 37, 1963, 22-26.

Method

Observers. The *O*s of this study were 37 men with at least 2 yr. of musical training who volunteered from an introductory course in psychology at Pennsylvania State University. All were between 18 and 24 yr. of age, and were selected from a large group of 60 volunteers on the basis of three criteria: (1) a life history of exposure to minimal noises as determined by Corso's Hearing Survey Questionary;[11] (2) a score above the 75th percentile in Seashore's test of pitch; and (3) a hearing loss of less than 5 db. from 250–8000 ~ on an audiogram of Békésy's type. From the group of 37 *O*s, the five obtaining the highest scores on the Seashore test were chosen to serve in the study dealing with ultrasonic pitch-discrimination, *i.e.* at 57,000 ~ .

The remaining 32 *O*s were assigned at random to the other four groups: Each

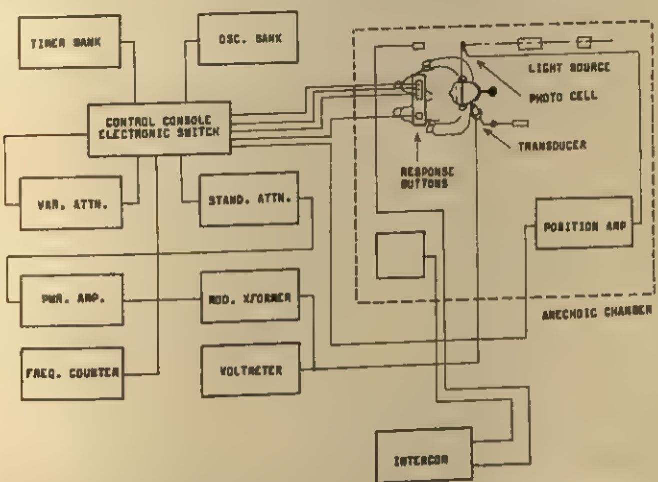

Fig. 1. Block Diagram of the Method of Bone-Conduction

group of 8 *O*s received practice at 2000 ~ and was tested at one other specified frequency by both air- and bone-conduction. Not all *O*s selected for the study were able to complete it satisfactorily due to illness or other uncontrollable factors: 3 from Group 2, 2 from Group 3, 2 from Group 4, and 1 from Group 5.

Apparatus. Fig. 1 is a block diagram of the apparatus for bone-conduction. Tones were generated by a bank of seven oscillators. One delivered the standard tone and the other six provided the comparison-tones. The signals from the oscillators were fed into an electronic switch which permitted *E* to present tones manually, or to present automatically a standard tone followed by one of the comparison-tones at a given time-interval. The standard-tones were each presented for 1.5 sec., separated by a silent interval of 0.5 sec. The interval between pairs of tones was 2.5 sec.

A General Ratio attenuator (Model 1450) was installed in each channel. The

[11] A detailed description of this questionary is available in Corso, Age and sex differences in pure-tone thresholds. *Arch. Otolaryngol.*, 77, 1963, 385-405.

output of the attenuators was then fed into an Eico power amplifier (HF-60). Since the transducer required 200v. for maximal output, a modulation transformer was employed to match the impedance of the amplifier to the relatively higher impedance of the transducer. A Hewlett-Packard voltmeter (Model 400 D) was connected across the input to the transducer to monitor the voltage.

The transducer was constructed from a barium titanate crystal, loaded at each end by a piece of machined aluminum. The response-curve of the transducer is relatively flat from 1000–16,000 ~ , but from 16,000–40,000 ~ its efficiency increases and the curve rises. From 40,000–100,000 ~ there are numerous resonant peaks where the efficiency is high and the device is free from measurable distortion. Technical details of the design of the transducer and its response characteristics have been presented elsewhere.[12]

O was seated in an anechoic chamber in a chair specially designed to provide maximal comfort and minimal movement when tilted approximately 20° to the rear. His head and neck were supported in a padded rest.

The transducer was attached to a free-swinging arm supported from the floor. Thus, *O*'s head was in contact only with the point of the transducer. The force applied by the transducer against *O*'s head was regulated by a set of weights attached to the transducer via a system of cords and pulleys. The force was maintained between 200 and 300 gr. following slight movements of *O*'s head.

A device signaling 'tilt' was used to monitor *O*'s head movements. This consisted of a metal rod which protruded 6 in. from *O*'s head on the side opposite to that being stimulated. Attached to this rod was a small photocell. A narrow beam of light coming from behind *O* was focused on the photocell. Whenever *O* moved more than a few degrees in any direction during the testing-period, the beam of light would shift from the cell and close a relay. This activiated a 'tilt'-light on both *O*'s and *E*'s instrument-panel and testing was suspended until *O* returned to his former position.

O was provided with a response-panel which contained three noise-free buttons. He was also provided with a light which flashed immediately prior to the onset of the standard stimulus. He could elect to use this signal or to turn it off.

Calibration of the transducer by bone-conducting was accomplished by placing the conical tip of the transducer in contact with the surface of a tank of water with a volume of approximately 3 cu. ft. A probe microphone, previously calibrated in the laboratory by the free-field method, was placed 2 mm. below the tip of the transducer. The distance selected was such as to minimize the effects of the standing wave, Calibration proceeded in 1000 ~ steps through the frequency range of 1000 to 100,000 ~ . The total distortion of the transducer at maximal output was less than 2.5% as determined by using a General Radio wave-analyzer, Model 736-A and a standardized procedure as described by Seely.[13]

Fig. 2 is a block diagram of the apparatus used in air-conduction. In most respects, it is quite similar to that for bone-conduction. The output of the console passed into a 40-w. power amplifier which had a frequency response of ±0.5 db. up to 80,000 ~ . At the rated output, harmonic distortion was less than 0.5%. A high-pass filter attenuated all components below 1,000 ~ at 12 db. per octave. The

---

[12] Corso and Oda, *op. cit.*, 22-27.
[13] Samuel Seely, *Electron-tube Circuits*, 1956.

voltage presented to the voice-coils of the transducer was monitored with a Ballentine Voltmeter (Model 300).

Two transducers were used, a horn-type dynamic speaker, (Electrovoice T350) used from 5000–20,000 ~ and an Acoustic Research speaker system (Model AR-2), used from 1000–5000 ~ . The two transducers were connected to the amplifier by a crossover network. To insure that the two units acted more or less as a point source of sound, the distances between *O*'s ear and the transducer was set at 30 cm. Measurements showed that sound-pressure followed the inverse-square law beyond this distance.

The acoustical components were calibrated using a Western Electric 640-AA

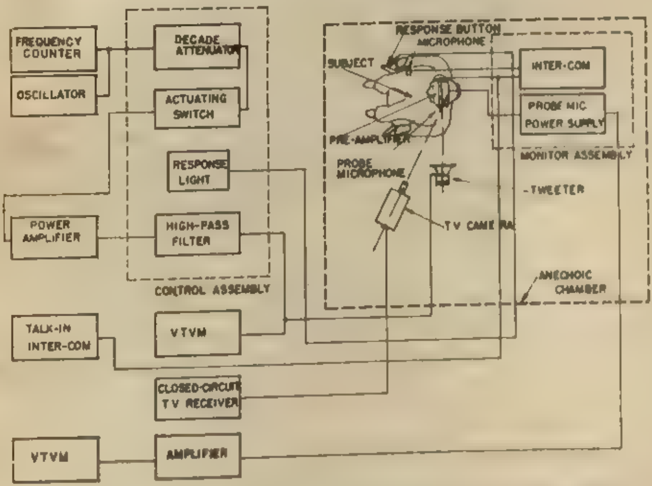

FIG. 2. BLOCK DIAGRAM OF THE METHOD OF AIR-CONDUCTION

condenser microphone. The characteristics of frequency-response, angular radiation-patterns, and harmonic distortion have been published elsewhere.[14]

Due to the extreme directionality of tones of high frequency, it was necessary to monitor the level of sound-pressure at the meatus during all experimental tests. To accomplish this, a probe tube was constructed which could measure the sound-field without significantly disturbing it. This system is also fully described elsewhere.[15] The probe microphone was set directly in front of the external auditory canal. The components of the system monitoring calibration are included in Fig 2. This arrangement permitted *E* to monitor instantly the sound-pressure of any stimulus sent over the system to *O*.

O was provided with the same response-system and tilt-indicator as used in testing bone-conduction. Since *O* was tested monaurally, however, he wore only one-half of the ear muff set in the tests by air-conduction.

Procedure. The experimental plan of the study was such that for Groups 1–4,

[14] Corso and Oda, *op. cit.*, 6-17.
[15] Corso and Oda, *op. cit.*, 17-21.

half of the *O*s in each group completed the tests with bone-conduction first and half with air-conduction first. The *O*s of all the groups served four times; two practice and two test sessions, approximately 1 hr. each. In the first practice session absolute threshold-determinations were made for air- and bone-conduction at 2000 ~ and at the frequency designated for the group to which *O* had been assigned. Each *O* then made at least 30 judgments of pitch for each of 6 comparison-stimuli (3 higher and 3 lower than the standard) by air- and bone-conduction. In the second practice session, all the *O*s made 50 judgments of pitch for each of six comparison-stimuli with a standard of 2000 ~ . Toward the end of this session, trials were given at the higher frequency for his group to establish for each *O* the value of the upper and lower comparison-stimuli which yielded 90% correct responses. The interval between these stimuli was then divided to yield six values to be used in the test-sessions, with the values spaced equally three above and three below the standard.

In the first test-session, the six predetermined comparison-stimuli for each *O* were presented 50 times each in random order. The second test-session was similar to the first, except that the alternate route of transmission, air or bone, was employed according to the counterbalanced plan. Throughout all sessions, the judgments of pitch were made at a 20-phon level of loudness and only judgments of 'higher' and 'lower' were permitted.

For Group 5, tested with a standard of 57,000 ~ , the procedure was somewhat different. At the first session, thresholds for bone-conduction were obtained at 2,000, 14,000; 16,000, 57,000, 64,000, and 94,000 ~ . In addition, trials of pitch-discrimination were given at 2000 ~ following the procedure already described. The second session was a replication of the first. In the third session, training in discrimination was started at 57,000 ~ with comparison-stimuli of 14,000, 16,000, 64,000 and 94,000 ~ . During this session, if *O* showed little or no discrimination, *E* would present the comparison-stimuli and indicate the correct judgment. In the fourth and fifth sessions, the same comparison-stimuli were employed, but *E* did not provide any guidance. Fifty judgments were made for each comparison-stimulus at a 20-phon level as determined from equal contours of loudness obtained for each *O*.

Results. The data for each *O* were initially treated separately to provide difference-limens for pitch at the frequencies tested in the five groups. The proportions judged higher were converted into z-values and a straight line was fitted to the data-points. The difference-limen was taken to be the probable error of the distribution. Individual *DL*s obtained at each frequency were then averaged to provide a measure of mean performance for each group.

Fig. 3 shows the mean *DL* as a function of frequency for both air- and bone-conduction. From 4000 ~ upward, there is a sharp rise in the magnitude of the *DL* for both methods of transmission. The curves also show that the *DL*s for air-conduction are larger than those for bone-conduction. The same data are plotted in Fig. 4 in terms of the *relative DL* for pitch as a function of frequency. Although the relative *DL* seems to remain

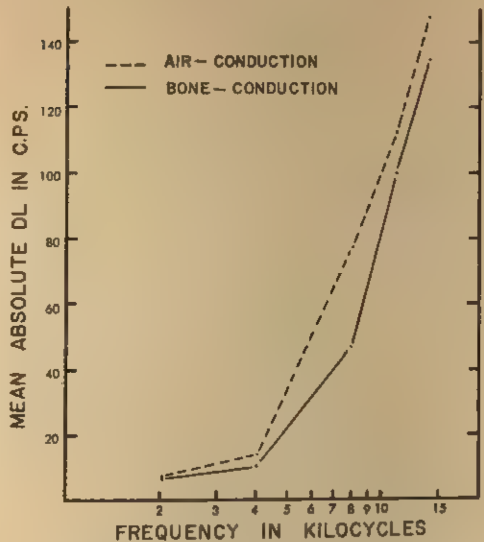

FIG. 3. ABSOLUTE DL FOR PITCH AS A FUNCTION OF FREQUENCY FOR BOTH AIR- AND BONE-CONDUCTION

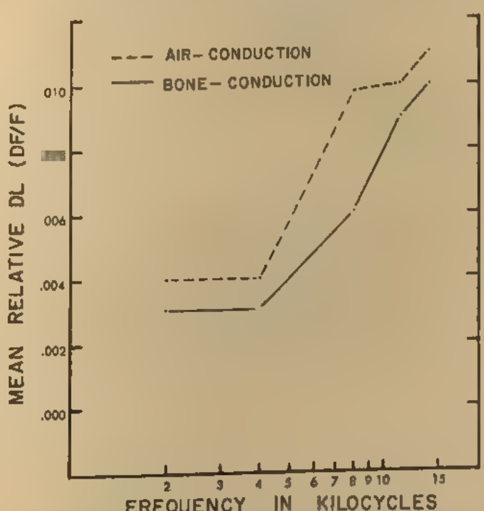

FIG. 4. RELATIVE DL FOR PITCH (DF/F) AS A FUNCTION OF FREQUENCY FOR BOTH AIR- AND BONE-CONDUCTION

constant up to 4000 ~ for both modes of transmission, the curves rise sharply thereafter with the maximal discrepancy between the curves occurring at 8000 ~.

Table I contains the mean absolute *DL*, from which Fig. 3 is plotted, and the standard error of the mean for air- and bone-conduction at each frequency tested above 2000 ~. As the mean *DL* increases, the standard error also increases. A Friedman two-way analysis of variance was employed to assess the reliability of the observed differences in the mean *DL*

TABLE I

MEAN ABSOLUTE DL FOR PITCH AND THE STANDARD ERROR OF THE MEAN FOR FOUR TESTED FREQUENCIES, IN CYCLES PER SECOND

Frequency	Number of *Os*	Air Conduction Mean *DL*	Air Conduction *SE*	Bone Conduction Mean *DL*	Bone Conduction *SE*
4,000	8	14.4	2.1	10.6	1.1
8,000	5	78.2	5.0	47.2	10.3
11,300	6	112.7	13.2	100.6	11.4
14,000	6	148.5	46.1	135.0	46.4

TABLE II

JUDGMENTS OF PITCH-DISCRIMINATION WITH A 57,000 ~ STANDARD

Each entry indicates the percentage judged higher in pitch.

O	Comparison stimuli in (cycles) 14,000	16,000	64,000	94,000
GT	43	87	50	97
AX	20	80	67	87
SK	30	60	87	77
BB	87	70	80	70
Mean %	45	74	71	83

for air- and bone-conduction.[16] The value of the statistic χ_r^2 was significant beyond the 0.05 level.

Table II shows the results obtained in the discrimination of the 57,000 ~ tone. The table shows the percentage of times the *Os* indicated that the comparison tones were higher than the standard. The 16,000 ~ tone was judged higher than the 57,000 ~ tone more often than a 64,000 ~ tone; the 14,000 ~ was judged higher nearly at the chance level. It seems reasonable to conclude, therefore, that pitch-discrimination is absent in the perception of ultrasonic frequencies presented via bone conduction.

Discussion. The data of the present study indicate that the form of the psychophysical function for pitch-discrimination is similar for both air- and bone-conduction. The *DLs* for bone-conduction are, however, significantly smaller than those for

[16] Sidney Siegel, *Nonparametric Statistics*, 1956, 166-172.

air-conduction. An explanation of this finding is suggested by the work of Shower and Biddulph.[17] They found that at equal sensation-levels, pitch-discrimination by bone-conduction is more sensitive than by monaural-air-conduction, but that bone-conduction was quite similar to binaural air-conduction. Thus, it appears that the differences in the two psychophysical functions are not attributable directly to the mode of transmission, but to loudness differences resulting from the stimulation of one ear versus two. When acoustic stimuli are presented via bone-conduction, the entire skull is placed into vibration and it is quite possible that the auditory mechanisms of the two ears are activated, even though a single transducer is applied to the region of one ear.

Pickler and Harris suggest that the discrepancy can be explained in terms of a loudness-difference between monaural and binaural stimulation, with the binaural condition appearing louder.[18] These investigators have found that at the same level of loudness, *i.e.* with the binaural channel (bone) matched in loudness to a monaural channel (air), the pitch *DLs* are essentially similar for air- and bone-conduction. For the binaural unmatched condition (as in the bone-conduction of the present study), the pitch *DLs* at 1000 ~ were approximately 0.5–1.0 ~ more sensitive than monaural air-conduction. In the present study and in the Shower and Biddulph study, however, the differences obtained between monaural air-conduction and unmatched bone conduction are considerably larger. The differences of Shower and Biddulph (all in favor of bone-conduction) range from 1 ~ at 1000 ~ to 24 ~ at 11,700 ~ ; in the present study, they range from 0.7 ~ at 2000 ~ to 13.5 ~ at 14,000 ~ . While the hypothesis of loudness-differential seems reasonable in the region below 4000 ~ , its adequacy at the higher frequencies seems doubtful in view of the large differences in *DLs*.

An alternative explanation for the air-bone difference lies in the nature of stimulation by bone-conduction. Several investigators have shown that there are two main types of stimulation by bone-conduction: translational and compressional.[19] Essentially, the difference between them is that in translational stimulation, the stimulation is provided in part by the relative movements of the ossicular chain and the cochlea. In compressional stimulation, the sides of the cochlea are compressed such that compensatory displacements of the basilar partition are produced with little or no contribution from the ossicular chain. Von Békésy has shown that from approximately 1,800 ~ upward, stimulation is primarily of the compressional type. Thus, if the influence of the ossicular chain were to be removed, there would be a loss in sensitivity where translational stimulation functions and no loss where compressional stimulation functions. Direct evidence on this has been reported by Smith, who found that when the ossicular chain is immobilized in the cat's ear, there is a bone-conduction loss in the low and middle frequencies, indicating that stimulation is of the translational type.[20] Above 2500 ~ , Legioux and Tarab have found that fixa-

[17] E. G. Shower and Rulon Biddulph, Differential pitch sensitivity of the ear, *J. acous. Soc. Amer.*, 3, 1931, 275-287.

[18] A. G. Pickler and J. D. Harris, Channels of perception in pitch discrimination, *ibid.*, 27, 1955, 124-131.

[19] Ernst Barany, A contribution to the physiology of bone conduction, *Acta Otolaryngol.*, 1938, suppl. no. 26, 233; Von Békésy, Vibration of the head in a sound field and its role in hearing by bone conduction, *J. acous. Soc. Amer.*, 20, 1948, 749-760.

[20] K. R Smith, Bone conduction during experimental fixation of the stapes, *J. exp. Psychol.*, 33, 1943, 96-107.

tion of the ossicular chain does not decrease the magnitude of the cochlear-microphonic potential in the guinea pig, indicating that stimulation is of the compressional type.[21] Since the frequencies used in the present study were at and above 2,000 ~ , it may be presumed that stimulation by bone-conduction was primarily compressional with no contribution from the ossicular chain. Inasmuch as the ossicular chain was functioning in air-conduction, it would appear, however, that the action of the middle ear may impose a limiting factor on the sharpness of the discrimination pitch and that the finest resolving power of the human ear may be found in bone-conduction.

The possibility of pitch-perception in the ultrasonic region has not been supported by the findings of the present study. The data at 57,000 ~ show no evidence for pitch-discrimination; however, the region beween 20,000 ~ and 57,000 ~ remains to be investigated.

Summary

The present study was performed to investigate pitch-discrimination for sonic and ultrasonic frequencies presented by monaural air- and by bone-conduction. Five groups of *O*s with normal hearing and high pitch-ability were tested at six frequencies from 2000 to 57,000 ~ at a loudness-level of 20 phons. The results obtained by the method of constant stimulus-differences indicate that: (1) at 2000 ~ the difference-limens for pitch discrimination are approximately equal for air- and bone-conduction, but from 4000–14,000 ~ the difference-limens are significantly smaller for bone-conduction; and (2) pitch-discrimination is absent for bone-conducted tones in the ultrasonic region. It is concluded that with respect to pitch-discrimination there is no functional difference between the two modes of acoustic transmission, but the apparent superiority of discrimination by bone-conduction suggests that there may be a loudness-difference between tones presented via air- and bone-conduction or that the role of the ossicular chain in air-conduction may impose a limiting factor on pitch-discrimination.

---

[21] J. P. Legioux and S. Tarab, Experimental study of bone conduction in ears with mechanical impairment of the ossicles, *J. acous. Soc. Amer.*, 31, 1959, 1453-1457.

SHORT-TERM MEMORY FOR PHONEMICALLY SIMILAR LISTS

By Wayne A. Wickelgren, Massachusetts Institute of Technology

A number of recent studies indicate that short-term memory for letters, digits, and words uses an auditory or speech-motor code. Conrad and Wickelgren have shown that errors in short-term recall of verbal lists tend to have a vowel or consonant phoneme in common with the correct item.[1] Perceptual errors were eliminated from these data by slow visual presentation or scoring of only items copied correctly during presentation. Proactive and retroactive interference in short-term recall are greater for interference lists consisting of letters that have a vowel phoneme in common with the correct letter(s) in the original list than for interference letters that have no phoneme in common with the correct letter(s).[2] Finally, lists of letters that are often confused with each other in auditory recognition are more difficult to recall than lists of letters that are rarely confused with each other in auditory recognition.[3] The former letters tend to have phonemes in common, while the latter letters tend to have no common phonemes.

If short-term memory is associative and verbal items are coded in short-term memory as sequences of phonemes, then the topological structure of the associations should be very different for lists of phonemically similar items than for lists of phonemically different items. Consider the list DZGBP as opposed to the list DQJYF. The first list consists of letters, all of which contain the vowel phoneme 'ē'; the second list has no cases of letters with common phonemes. Fig. 1 describes what one phonemic-associative theory of short-term memory would predict concerning the direct phoneme-to-phoneme associations formed during presentation of each of these two lists. The only associations shown in Fig. 1 are the forward and backward intra-item associations and the forward 'direct' associations between the vowels

* Received for publication January 8, 1965. This work was supported primarily by grant, MH 08890-01, from the National Institute of Mental Health, U.S. Public Health Service. Further aid was received from grant, MH 05673-03, to Hans-Lukas Teuber.

[1] R. Conrad, Acoustic confusions in immediate memory, *Brit. J. Psychol.*, 55, 1964, 75-84; W. A. Wickelgren, Acoustic similarity and intrusion errors in short-term memory, *J. exp. Psychol.*, 70, 1965, 102-108.

[2] Wickelgren, Phonemic similarity and interference in short-term memory for single letters, *J. exp. Psychol.* in press; Acoustic similarity and retroactive interference in short-term memory, *J. verb. Learn. verb. Behav.*, 4, 1965, 53-61.

[3] R. Conrad, and A. J. Hull, Information, acoustic confusion and memory span, *Brit. J. Psychol.*, 55, 1964, 429-432.

and consonants of adjacent items. Many important associations have been omitted to reduce the complexity of the diagrams. Associations between serial position concepts (such as 'beginning,' 'middle,' and 'end') and phonemes have been omitted because there is no reason to suppose the structure of these associations is different in any important way for the two lists. Backward and remote inter-item associations are presumed to be weaker, on the average, than the more direct associations shown in Fig. 1, but they should be considered as competing associations which are some-

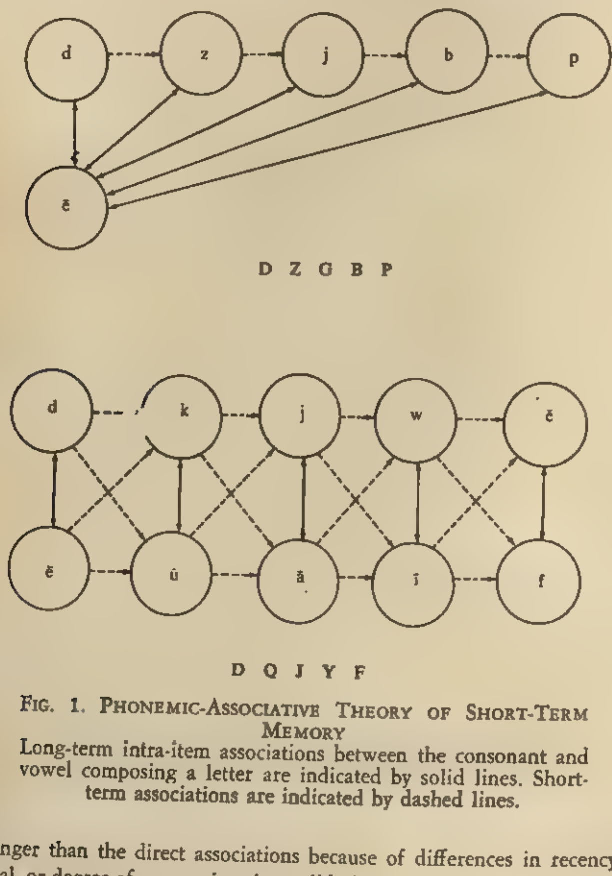

FIG. 1. PHONEMIC-ASSOCIATIVE THEORY OF SHORT-TERM MEMORY
Long-term intra-item associations between the consonant and vowel composing a letter are indicated by solid lines. Short-term associations are indicated by dashed lines.

times stronger than the direct associations because of differences in recency, amount of rehearsal, or degree of unconscious 'consolidation.'

The difference in topological structure of the associative network for the two lists is obvious. The memory-trace for ordered recall is much better replicated in the case of phonemically different lists than in the case of phonemically similar lists. In the simplified model described in Fig. 1, there are four associations between the representatives of adjacent letters to carry the trace for ordered recall of the phonemically different list and only one association (between adjacent consonants) to carry the trace for ordered recall of the phonemically similar list. Free recall of

the phonemically similar list should not be at such a disadvantage. In fact, one might well argue that free recall of items should be better for phonemically similar lists since the 'ē' phoneme is certain to be recalled and direct associations exist from the representative of the 'ē' phoneme to the representatives of all the consonant phonemes in the list. If competition of response blocks recall of the consonants whose associations to the representative of the 'ē' phoneme are weakest, then free recall of phonemically similar lists might be poorer than free recall of phonemically different lists. But if competition of response does not prevent recall of the consonants whose associations to the representative of the 'ē' phoneme are weakest, then we might expect more items to be recalled from phonemically similar lists, though often in the wrong positions.

Conrad and Hull's study of short-term memory for acoustically similar lists did not analyze recall separately for recall of items and recall of position.[4] If the phonemic-associative theory just described is correct, then the superior ordered recall of phonemically different items should be reflected primarily in better recall of the position of an item. The number of items recalled, irrespective of position, should be less affected by phonemic similarity and might show a reversed effect.

The following two experiments were designed to test the above predictions using lists of letters and lists of consonant-vowel diagrams (*e.g.* fā, tā, kā, gā, dā, vā, bā, or fā, lē, mī, vō, nōō, zī, dē). In the latter case, only the initial consonant of each diagram was to be recalled, so there was no difference in the number of different phonemes to be recalled. Also, with diagrams there can be no cognitive strategies based on long-term associations between the consonant and vowel phonemes composing a letter.

Experiment I

Procedure. Presentation was auditory, and the entire experiment was recorded on tape. Subjects listened to a 'ready' signal followed after 1 sec. by a list of nine letters presented at the rate of three letters per sec., followed immediately by an attempt to recall the letters in the correct order (by filling-in nine boxes). Twenty seconds were allotted to recall, followed immediately by the next trial.

Design. Eight types of lists were used in the experiment. Each *S* was given every type of list, randomly ordered in 12 blocks of eight, for a total of 96 lists. All lists were 9 letters long, but ony 4 or 6 of these letters were different. Lists, however, were so constructed that there were no runs of two adjacent identical letters. The types of lists were as follows: (Pure ā—4) Random list of 'ā' letters (A, H, J, K) using each of the four letters at least once. (Pure ĕ—4) Random list using four 'ĕ' letters (selected from F, L, M, N, S, X), with the same repetition structure as the ā—4 list in the same block of eight lists. For example, if the ā—4 list was HAHJKAHJA, then the ĕ—4 list might be MSMLFSMLS. (Pure ē—4) Random list using four 'ē' letters (B, C, D, E, G, P, T, V, Z), with the same repetition structure as the ā—4 list in the same block. (Mixed ā, ĕ, ē—4) Random list using four different letters, at least one of each of the three types, with the

[4] Conrad and Hull, *op. cit.*, 429-432.

same repetition structure as the ā—4 list in the same block. (Pure ĕ—6) Random list using the six different 'ĕ' letters. (Pure ē—6) Random list using six 'ē' letters, with the same repetition structure as the ĕ—6 list in the block. (Pure ā + ū—6) Random list using six different letters drawn from the 'ā' letters and the 'ū' letters (Q, U, W), with the same repetition structure as the ĕ—6 list in the same block. (Mixed ĕ, ē, ā + ū—6) Random list using six different letters, two from each of the three types, with the same repetition structure as the ĕ—6 list in the block. The purpose of this design is to permit comparison of pure and mixed lists without confounding by differences in the difficulty of different letters or differences in the repetition structure of the lists.

Subjects. The *S*s were 31 undergraduates in psychology who participated in the experiment as a part of their course requirements. They were used in two approximately equal groups.

Results. Each *S*'s data for each condition were analyzed for ordered recall, item-recall, and position-recall for items (letters). An *S*'s report of an item is correct by an ordered recall criterion if and only if the correct item is recalled in the correct position. An *S*'s report of an item is correct by an item-recall criterion if and only if it appears anywhere in his report of the list in question. The ordered and item-recall error rates for a condition for an *S* are the ratio of the ordered or item-recall errors, respectively, to the total number of opportunities for error. Position-recall, independent of item-recall, is obtained by subtracting the item-recall errors from the ordered recall errors, reducing the total number of possible errors by the number of item-recall errors when computing the error rate for position-recall. Item-recall and position-recall are statistically independent, and ordered recall reflects the combined operation of the recall of items and the recall of the correct position for items. Ordered recall is not statistically independent of item-recall, and therefore, is not as suitable for the assessment of the recall of order as is position-recall. Position-recall is statistically independent of item-recall because only the items that are correct by an item-recall criterion are scored for position-recall.

When any two conditions are compared by any of the three measures of recall, the error rate in each condition for each *S* is computed and one is subtracted from the other to give a difference score favoring one condition or the other. Superiority of one condition over the other is determined by applying the Wilcoxon Matched-Pairs Signed-Ranks test to the 31 difference scores.

The error-rates for ordered recall, item-recall, and position-recall in each type of list are reported in Table I. The error-rates for ordered recall were significantly greater in pure lists than in mixed lists, replicating

the results obtained by Conrad and Hull.[5] Exactly as predicted by the phonemic-associative theory, this difference resulted entirely from significantly greater error in the recall of position in the pure lists. Errors in item-recall showed small conflicting differences and were not significantly different overall.

Lists with only four different letters (and therefore more repetition of letters) were remembered better than lists with six different letters; the difference between reflected primarily in item-recall rather than position-recall. Error rates were significantly lower in ĕ—4 and ē—4

TABLE I

Error Rates (%) in Experiment I

List type	Ordered-recall	Item-recall	Position-recall
Pure ā—4	38.4	18.2	24.7
Pure ĕ—4	48.4	23.2	32.9
Pure ē—4	53.2	26.7	36.2
Total pure—4	46.7 †	22.7 *	31.0 †
Mixed ā, ĕ, ē—4	39.7	24.6	20.1
Pure ĕ—6	53.1	27.4	35.4
Pure ē—6	56.2	33.0	34.6
Pure ā+ū—6	44.3	25.2	25.5
Total pure—6	51.2 †	28.5 *	31.7 †
Mixed ĕ, ē, ā+ū—6	40.8	26.3	19.7
Total pure	49.0 †	25.6	31.4 †
Total mixed	40.2	25.5	19.9

* $p<0.05$ † $p<0.001$.

than in ĕ—6 and ē—6 in ordered recall ($p < 0.001$) and item-recall ($p < 0.001$), but there was no significant difference in position-recall.

Finally, there were some large differences between different types of letters; 'ā' letters were remembered better than 'ĕ' letters which were remembered better than 'ē' letters. With one insignificant exception, the relationship between the types of letters was the same for ordered recall, item-recall, and position-recall. Explanation of these differences between types of letters is beyond the scope of the study.

Experiment II

Procedure. Presentation was auditory and recorded on tape. The *Ss* listened to a 'ready' signal, followed after 1 sec. by seven consonant-vowel (CV) diagrams presented at the rate of two diagrams per sec., followed immediately by an attempt to recall the consonants of the diagrams in the correct order (by filling-in seven boxes). Sixteen seconds were allotted to recall, followed immediately by the next trial.

[5] Conrad and Hull, *op. cit.*, 429-432.

Design. Six types of lists were used in the experiment. Each *S* was given every type of list, randomly ordered in 10 blocks of 10 lists, for a total of 100 lists. The five 'pure' types of lists occurred once each in a block of 10 and the single 'mixed' type of list occurred 5 times in a block of 10. The 7 different consonants in each list were selected randomly from the following set: b, d, f, g (as in 'go'), k, m, n, p, s, t, v, z. The 6 types of lists were as follows: (Pure ā) All 7 consonants were followed by 'ā' (*e.g.* sā, gā, vā, bā, nā, tā, fā). (Pure ē) All 7 consonants were followed by 'ē'. (Pure ī) All 7 consonants were followed by 'ī'. (Pure ō) All 7 consonants were followed by 'ō'. (Pure ōō) All 7 consonants were followed by 'ōō'. (Mixed) The 7 consonants were followed by one of the 5 vowels, with 3 of the 5 vowels being used once and the other two vowels being used twice in each list of diagrams (*e.g.* fōō, nē, bā, vō, gā, mī, zē).

Subjects. The *S*s were 28 undergraduates in psychology who participated in the experiment as a part of their course requirements. They were used in two approximately equal groups.

Results. The error rates for ordered recall, item-recall, and position-recall in each type of list in Experiment II are reported in Table II. The

TABLE II

ERROR RATES (%) IN EXPERIMENT II

List type	Ordered-recall	Item-recall	Position-recall
Pure ā	49.6	27.3	30.7
Pure ē	50.7	40.6	29.0
Pure ī	53.8	31.0	33.0
Pure ō	50.4	27.9	31.2
Pure ōō	48.2	27.7	28.3
Total pure	50.5*	28.9†	30.4†
Mixed	47.8	31.9	23.3

* $p<0.05$ † $p<0.001$.

over-all ordered recall of pure vs. mixed lists showed a significantly higher error rate in the pure lists ($p < 0.05$). Again, this difference in ordered recall resulted from a significantly higher error rate in position-recall for pure lists ($p < 0.001$), which masked a smaller, but highly significant, difference in the opposite direction for item-recall ($p < 0.001$).

Again there were some differences between different types of pure lists, but the differences were smaller than those in Experiment I. By and large, any difference of 3% or larger in Table II indicates a significant difference at the 0.05 level using the Wilcoxon Matched-Pairs Signed-Ranks test.

Discussion. Both of the present experiments support the phonemic-associative theory of verbal short-term memory. Phonemically similar lists are more difficult to recall in order than phonemically different lists primarily because the items from phonemically similar lists are much more likely to be recalled in the wrong positions than are the items from

phonemically different lists. The number of items recalled correctly, irrespective of position, generally shows a small advantage in favor of the phonemically similar lists. The most pronounced dissociation of item-recall and position-recall was obtained in short-term recall of the initial consonant of diagrams. Thus, it is not possible to argue that long-term associations (such as those involved in our knowledge of the phonemic characteristics of letters in the alphabet) are responsible for the superior item recall of phonemically similar lists.

It seems difficult to account for the present findings and previous findings on phonemic similarity without assuming that items are coded in short-term memory as sequences of phonemes, at some level of analysis. Naturally, we cannot say from these data whether the phoneme is the ultimate unit of coding in short-term memory. The results of Miller and Miller and Nicely indicate that in auditory recognition of vowels and consonants there is a level of analysis beyond the phonemic level, namely distinctive feature analysis.[6] It may be that the distinctive feature is also a more basic unit of coding in short-term memory, and perhaps there are levels of analysis even more basic than distinctive features. The present findings indicate only that, whatever the most basic units are, these basic units combine to represent phonemes, which in turn combine to represent letters, digits, words, etc.

It also seems difficult to account for the present findings with a non-associative theory of short-term memory. Non-associative theories of short-term memory are those in which an ordered set of cells (boxes, locations, registers, etc.) are set aside as temporary ('buffer') storage and any item can be encoded into any cell. Ordered recall is possible because a list of items is generally or always stored in the cells in a fixed order and generally or always retrieved from the cells in the same order. To account for the present findings a non-associative theory must make additional assumptions like the following: (1) The correspondence between storage order and retrieval order is poorer for lists with greater phonemic similarity. (2) The rate of trace decay for an item within a cell is slower when phonemically similar items are stored in other cells than when phonemically different items are stored in these other cells. Both assumptions are completely *ad hoc,* and the second assumption seems incompatible with the greater retroactive interference in item-recall produced by phonemically similar interpolated items.[7]

[6] G. A. Miller, The perception of speech, in *For Roman Jakobson,* 1956, 353-359; G. A. Miller, and P. E. Nicely, An analysis of perceptual confusions among some English consonants, *J. acoust. Soc. Amer.,* 27, 1955, 338-352.

[7] Wickelgren, *op. cit., J. verb. Learn. verb. Behav.,* 53-61; Phonemic similarity and interference in short-term memory for single letters, *J. exp. Psychol.,* in press.

An associative theory of short-term memory requires no additional assumptions to account for the greater interference of similar interpolated items and the greater item-recall in similar lists. Whether the similar items are in the list to be recalled or in an interpolated list, they produce competing responses to the common phonemes. When the comparison is between similar and different *interpolated* items, these competing responses are not correct by an item-recall criterion since they are not in the original list. When the comparison is between similar and different items in the original list, the competing responses are all correct by an item-recall criterion. The present findings indicate that in short-term memory, competing associations to a common phoneme have little tendency to block the weaker associations, so long as one is allowed to recall as many items as there are competing responses. Whatever blocking of weaker associations does exist, its effect on item-recall is generally less than the advantage of having a cue that is associated to all the correct responses.

Summary

In the first experiment 31 *Ss* attempted ordered recall of two types of 9 letter lists: phonemically similar lists in which all letters had a common vowel phoneme (ā, ē, or ĕ) and phonemically different lists whose letters had no common phoneme. Ordered recall was poorer for similar lists ($p < 0.001$), but this resulted entirely from poorer recall of the position of similar letters ($p < 0.001$). Item-recall, by a free recall criterion, was not significantly different for the two types of lists. In the second experiment 28 *Ss* attempted ordered recall of the consonants only, from two types of lists of seven consonant-vowel diagrams: phonemically similar lists in which the vowel was identical for all seven diagrams (ā, ē, ī, ō, o͞o) and phonemically different lists whose seven vowels were a mixture of the above five vowels. Position-recall was significantly poorer for phonemically similar lists ($p < 0.001$), but item-recall was significantly better for similar lists ($p < 0.001$).

THE ROLE OF KNOWLEDGE IN DISTANCE-PERCEPTION

By Howard E. Gruber, New School for Social Research, and
Albert J. Dinnerstein, New York Medical College

This paper deals with the relation between perception and knowledge in the judgment of distance. Although both depend on the individual's contact with his environment, in the moment of experience, they are different; *seeing* does not feel like knowing.

In this study, perception was defined phenomenally: *O* was asked to describe "the way things look to you at the moment." The definition of knowledge, implicit in the operations by which knowledge was induced, included both subjective and objective criteria: the stimulus-objects were in fact at very different distances from *O*, he was convinced that this was the case, and his conviction was based on evidence. As a concept, knowledge must be differentiated from past experience. It has been argued, for example, that past experience influences the interpretation of depth-cues, but no one has yet argued that the perceiving individual necessarily knows of the existence of these cues. Knowledge is an aspect of consciousness; only under certain conditions does past experience produce it.

Although the effects of knowledge on perception long have been at issue in psychological theory, there have been few studies in which this variable was experimentally manipulated. In one study by Brunswik, the effects of knowledge on the Müller-Lyer illusion were investigated.[1] Henneman and MacLeod studied the effect of knowledge on brightness-constancy.[2] Many studies have dealt with the effects of known size on perceived distance. Those with positive results have been criticized on the ground that they really demonstrate the effect of relative size on perceived distance.[3] Negative results may not represent the failure of knowledge

* Received for publication December 1, 1964.

[1] Egon Brunswik, *Perception and the Representative Design of Psycohlogical Experiments*, 1956, 93-99.

[2] R. H. Henneman, A photometric study of the perception of object color, *Arch. Psychol.*, 179, 1935, 1-88; R. B. MacLeod, Brightness-constancy in unrecognized shadows, *J. exp. Psychol.*, 27, 1940, 1-22.

[3] C. B. Hochberg and J. E. Hochberg, Familiar size and the perception of depth, *J. Psychol.*, 34, 1952, 107-114; William Epstein, John Park, Albert Casey. The current status of the size-distance hypothesis, *Psychol. Bull.*, 58, 1961, 491-514; W. C. Gogel, *The Visual Perception of Size and Distance*, Report No. 62-15 from Civil Aeromedical Research Institute, 1962; J. E. Hochberg, Nativism and empiricism in perception, in Leo Postman (ed.) *Psychology in the Making*, 1962, 255-330.

to influence perception, but only the failure of *known size* to influence *perceived distance*. The present study deals with what may well be a more direct relationship—*the influence of known distance on perceived distance.*

Experiment I

The experimental situation provided room for conflict between known and perceived distance in such a way as to distinguish between perceived absolute distance (*O*'s impression of distance from self to object) and perceived relative distance (ratio of perceived absolute distances of two objects). Gogel has recently used "relative distance" to mean a *difference* between two absolute distances, not a ratio.[4] His use of the term "absolute distance" is the same as ours.

Apparatus. The stimuli, displayed in a long corridor, were two luminous squares subtending equal visual angles, produced by identical light-boxes with milk-glass panels and suitable masking arrangements. The right-hand square was 2 in. on a side and 24 ft. from the *O;* the left-hand square was 4 in. on a side and 48 ft. from *O*. The separation of the two squares was 3°30′ of visual angle between adjacent edges. In the dark, *O* saw only two luminous squares. When the corridor was illuminated, *O* saw the light-boxes mounted on standards, their points of contact with the floor of the corridor, and numerous objects scattered along the sides of the corridor.

Procedure. O was seated and a nose-rest so adjusted that both stimulus-squares were centered at eye-level. One eye was masked. The door to the darkened corridor was now opened, revealing the two squares, and *O* asked to describe their apparent distances. Then the corridor was illuminated and *O* again described what he saw. This procedure was repeated for 12 cycles of dark and light observations. The instructions were designed to elicit phenomenal judgments of distance, expressed in feet and inches. The *O*s were 16 summer-school students in psychology.

Results. Knowledge of distance had a statistically significant effect on perceived absolute distance, which was calculated as the sum of the judgments for the left and right squares, divided by 2 (Trial 1: $Md = 14.50$ ft., $AD = 8.10$; Trial 2: $Md = 23.25$ ft., $AD = 12.10$; $p < 0.01$, Wilcoxon's test for paired replicates).

Knowledge of distance did not have a statistically significant effect on perceived relative distance (Trial 1: $Md = 1.00$; $AD = 0.11$; Trial 2: $Md = 1.06$, $AD = 0.31$). Although the objective difference in distance of the two squares was 24 ft., the median reported perceived difference on Trial 2 (post-knowledge) was only 5.5 in. When the corridor was darkened, at the start of the second cycle, most *O*s reported with surprise that the depth in the previously illuminated scene faded rapidly. This

[4] Gogel, 1962, *op. cit.*

fading of depth, or mutual approach of the two squares, reached its limit in about 10 sec.

With several additional *O*s, it was found that alternating between binocular and monocular observation in the darkened corridor produced a dramatic shift in perceived distance, leaving the appearance of the two squares unaltered in other respects.[5] The next experiment makes use of this effect.

Experiment II

In Experiment II a 3 × 2 design was used. Half the *O*s were controls: they never saw the lighted corridor. The other half saw and explored the lighted corridor before their first encounter with the luminous squares in the darkened corridor. Each of the two groups was further subdivided into three groups in terms of the retinal-size ratio of the two squares. There were thus six groups in all, with nine *O*s per group, all college students.

Apparatus. The apparatus was the same as in the preceding experiment, except that three different masks were made for the right-hand light-box. They provided retinal-size ratios (near square: far square) of 0.5, 1.0, and 2.0. The right square was 24 ft. from *O* and the left square 48 ft. from *O*.

Training. Training was intended to insure that all *O*s would understand the distinction between knowledge and perception, and would give distance-judgments in perceptual terms. The essential parts of the instructions are given verbatim.

As you probably realize, things do not always *appear* exactly the way they are. For example, from a great distance or from the top of a tall building, people usually look very small. In this experiment we want you to report the *appearance* of certain objects, especially their distances.

Several examples of optical illusions were then presented to *O*: a perspective drawing in which depicted depth produces dramatic differences in the perceived size of objects, the Müller-Lyer figure, the Hering illusion, the Zöllner illusion, and the vertical-horizontal illusion. A ruler was offered to *O* with which to measure the figures, and *E* then continued:

As you can see, there are sometimes differences between the way things appear and the measurements obtained with a ruler. In the experiment we are going to do now, this may or may not be the case; your task is simply to describe your impressions—or the way things look to you at the moment.

After this preliminary period, the nose-rest was so adjusted that the centers of the two squares both would be at eye-level and their edges equidistant from the median plane during the actual judgments. For the control groups, a series of nine distance-judgments of the luminous squares followed immediately, with rest-periods between judgments equal to the time taken by the experimental group for the contacts with the corridor. The door to the corridor was closed during these rest-

[5] For simplicity, changes in perceived size are ignored throughout this study.

intervals. For the experimental groups, the following procedure was used before each trial: *E* accompanied *O* on a walk down the corridor, pointing out the light-boxes which provided the stimuli, and the chalked markings on the floor indicating various distances in feet from the point of observation. The *O* took his seat, placed his head in position, and observed the corridor *binocularly* in *full illumination.* The lights in the corridor were turned out, and *O* explored the corridor with a flashlight. Then *O* turned off the flashlight and observed the two luminous squares, still with binocular vision. Finally, *O* closed one eye, and, after a 5-sec. pause, made his judgment. This procedure was designed to produce as much continuity as possible between the knowledge-inducing situation and the judgmental

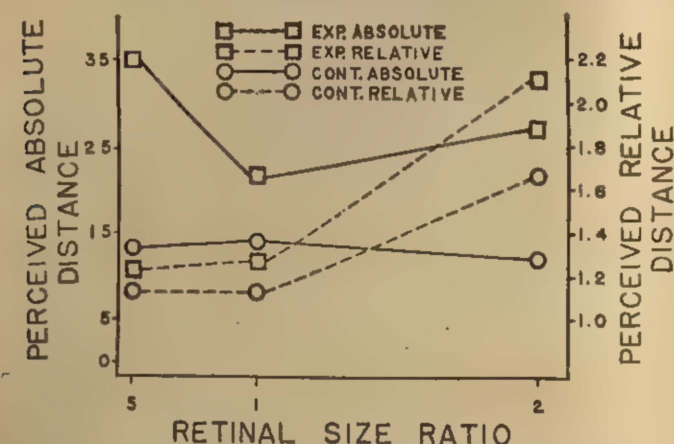

FIG. 1. JUDGMENTS OF ABSOLUTE AND OF RELATIVE DISTANCE MADE BY EXPERIMENTAL GROUPS HAVING KNOWLEDGE OF TRUE DISTANCES AND BY CONTROL GROUPS

situation. After *O* gave his distance-judgments, the lights were turned on and the knowledge-producing procedure was repeated as before, except that *E* did not accompany *O* on his walk down the corridor after the first trial.

Results. As in Experiment I, the data were analyzed with respect to perceived absolute distance and perceived relative distance. In Fig. 1, the results are given in terms of the actual judgments. Each plotted point represents the median for 9 *O*s, each median based on the mean of 9 trials for each *O*. For perceived absolute distance, *judgment* refers to the sum of the reported distances to left- and right-hand squares divided by two. For perceived relative distance, *judgment* refers to the ratio (perceived distance of left object/(perceived distance to right object).

Fig. 1 suggests that some residual distance-cues were present in the

stimulus-situation, since the control group's judgments yielded median ratios greater than one under all three stimulus-conditions. This fact does not vitiate the results to be presented below, since these cues were present equally for control and experimental groups.

The data were treated by non-parametric analyses of variance, both with judgment-scores and with discrepancy-scores.[6] Both analyses yielded the same results. In the analyses of relative-distance judgments, the effect of stimulus-size was significant ($p < 0.01$) but the effect of knowledge was not significant. The analyses of absolute-distance judgments showed the opposite result: knowledge was significant ($p < 0.01$) but stimulus-size was not significant. In no case was the interaction of knowledge and stimulus-size significant. The data also were examined for changes in judgment in the course of the nine trials.[7] No significant trend was found.

Discussion. The most provocative result of these studies is that knowledge markedly influenced the perception of absolute distance but had little effect on the perception of relative distance.

Did the changes in reports of absolute impressions of distance reflect actual changes in perception, or were they merely changes in the use of numerical scales? Several lines of evidence suggest the former conclusion. First, in Experiment I, when the corridor was illuminated for the first time, the *O*s were surprised and frequently remarked on the disparity between their impressions in the light and in the dark; furthermore, when the lights were turned off and the impression of distance again changed, the *O*s seemed equally surprised. Secondly, the fact that knowledge acted selectively argues against the proposal that the changes were merely nominal. If the *O*s were simply changing verbal reports to conform with knowledge, there was no particular reason to change with regard to absolute but not to relative distance. The *O*s' comments indicated a clearcut awareness of this selective character of the perceptual change. Finally, each *O* in Experiment II was interviewed briefly after the experiment. The most important question was, "Did you actually report the way the distances looked?" Of 27 *O*s in the experimental groups, all but one gave an unqualified affirmative answer. The task of giving phenomenal reports at variance with knowledge was surprisingly easy. In the experimental groups, *O*s were asked, "Did you find it difficult to report that the distances looked different than you knew them to be?" On a 3-point rating

[6] F. D. Sheffield, Comment on a distribution free factorial-design analysis, *Psychol. Bull.*, 54, 1957, 426-430.

[7] Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences*, 1956.

scale, 22 *O*s indicated no difficulty, 3 *O*s indicated slight difficulty, and 2 *O*s indicated great difficulty.

Perhaps most damaging to the view that knowledge determines the interpretation of depth-cues was the striking fact that a fully informed *O* could watch perceived depth fade when he simply closed one eye.[8] Knowledge did, however, affect perceived absolute distance, and this result conflicts with an extreme stimulus-determination-theory, *i.e.* the view that perception is *nothing but* a function of stimulation.

Regarding the distinction between effects of knowledge on perceived absolute and on perceived relative distance, there is no previous work. Two experiments by Gogel at least seem to demonstrate that perceived absolute distance and perceived relative distance can be distinguished experimentally in that they respond differently to variables of stimulation.[9] The findings of Gibson and Bergman, of Gibson, Bergman, and Purdy, and of Purdy and Gibson parallel ours:[10] training improved perception of absolute distance but did not effect perception of relative distance.

The present results have some relevance for the size-distance invariance hypothesis. Gruber has pointed out the need to select the appropriate measure of perceived distance to be correlated with perceived size in studies of this problem.[11] Several recent investigators have varied the type of size-judgment obtained in the hope of specifying the conditions under which the size-distance invariance hypothesis holds true.[12] Their results have been contradictory and inconclusive. The present study suggests the importance of examining a variety of measures of perceived dis-

---

[8] Did the use of objects close together in the field of view mask the effect of knowledge by introducing a countervailing tendency to see such objects as equidistant? W. C. Gogel and G. S. Harker (The effectiveness of size cues to relative distance as a function of lateral visual separation, *J. exp. Psychol.*, 50, 1955, 309-315) have shown that lateral proximity can mask the influence of relative size on perceived relative distance under some conditions of observation. The marked, significant effect of relative size on perceived relative distance in Experiment II indicates that lateral proximity was far less effective in this experiment than in Gogel and Harker's.

[9] W. C. Gogel, *op. cit.*

[10] E. J. Gibson and Richard Bergman, The effects of training on absolute estimation of distance over the ground, *J. exp. Psychol.*, 48, 1954, 473-482; E. J. Gibson, Richard Bergman, and Jean Purdy. The effect of prior training with a scale of distance on absolute and relative judgments of distance over ground, *ibid.*, 50, 1955, 97-105; J. Purdy and E. J. Gibson, Distance judgment by the method of fractionation, *idem*, 374-380.

[11] H. E. Gruber, The relation of perceived size to perceived distance, this JOURNAL, 67, 1954, 411-426.

[12] V. R. Carlson, Size-constancy judgments and perceptual compromise, *J. exp. Psychol.*, 63, 1962, 68-73; J. C. Baird, Retinal and issumed size cues as determinants of size and distance perception, *ibid.*, 66, 1963, 155-162; William Epstein, Attitudes of judgment and the size-distance invariance hypothesis, *idem*, 78-83.

tance in relation to a variety of measures of perceived size. If the invariance hypothesis holds true at all, it may only apply when the appropriate *combination* of measures is employed.

Summary

The distances of nonequidistant pairs of luminous squares in a dark field were judged by college students in two experiments. Control groups judged without knowledge and experimental groups with knowledge of the objective distances. In Experiment I, the stimuli were of the same retinal size. Knowledge of distance influenced perceived absolute distance but not perceived relative distance. In Experiment II, the retinal-size ratio of the stimuli was varied. Again, knowledge influenced perceived absolute distance but not perceived relative distance. Stimulus-size ratio influenced perceived relative distance but not perceived absolute distance. Viewed binocularly, the stimuli appeared nonequidistant; monocularly, they seemed equidistant or almost so. By opening and closing one eye, this phenomenal shift could be repeated indefinitely. These results show that knowledge has little influence on perceived relative distance.

EXPANSION AFTER-EFFECTS WITHOUT PERCEIVED CONTRACTION OF THE INSPECTION-FIGURE

By HERMAN H. SPITZ, Bordentown, New Jersey

The present experiments arose from a question concerning an illusory spiral illustrated by Boring, who adapted it from Fraser.[1] It is shown here as Stimulus A in Fig. 1. Since this is not a true spiral, but actually a series of concentric circles made up of slightly tilted, arched lines, what would occur were it spun in a clockwise direction?

In fact, when the figure is spun at 16 r.p.m., the illusory spirial appears to expand. When stopped, an after-effect of inward spiral movement is observed. To determine the contribution of the background to these effects, progressively more of the background was eliminated until there remained only a few circles made up of the short, inclined lines. At this point an unusual phenomenon was noted: when the figure was spun in a counter-clockwise direction no 'contraction' was produced, but when stopped a slight 'expansion' was observed, along with a strong clockwise movement. By reversing the tilt of the lines, straightening their arcs (see A, Fig. 2), spinning the figure in a clockwise direction, and using true, matched concentric circles as a test (T)-figure (see A, Fig. 3), a strong expansion after-effect could be produced, and without observed contraction of the spinning inspection (I)-figure. To our knowledge, this was the first time that movement after-effects have been observed without the prior perception of movement in the opposite direction.

Method: Subjects. These observations were then subjected to experimental verification, with 120 volunteer *Ss*, women students from Trenton State College.[2] (average age = 19.41 yr., $\sigma = 0.59$ yr.) Fifteen *Ss* were randomly placed in each of eight groups and individually tested. There were no significant age-differences between any of the groups.

Instructions. To determine whether the *Ss* perceived inward or outward movement of the *I*-figure, they were instructed to fixate the center of the spinning figure and to report exactly what they saw, and if anything changed. Each was asked to

---

* Received for publication February 1, 1965. From the Edward R. Johnstone Training and Research Center.

[1] E. G. Boring, The perception of objects, *Amer. J. Physics*, 14, 1946, 99-107; James Fraser, A new visual illusion of direction, *Brit. J. Psychol.*, 2, 1908, 307-320; Plates 1–9.

[2] The author is most grateful to Marshall P. Smith, Chairman of the Psychology Department at Trenton State College, for securing subjects and providing space for this study.

become a "reporting machine," reporting everything about the whole figure. They were also told that, after 1 min., another figure, made up of circles, would be brought into view, and that they were to keep their eyes on the center dot and report whether the circles were getting bigger, staying the same, or getting smaller. The latter part of these instructions was repeated 12 sec. prior to lowering the *T*-figure into view.

Materials and apparatus. The *I*-figures were disks, measuring 8 to 8½ in. in diameter, which fitted on a vertical fixed 8-in. phonograph turntable, so that clockwise spin was used throughout the experiment. The turntable protruded slightly from a background of homasote, painted flat black, which shielded *E* from view.

The *T*-figures were inserted in a 9½ × 10½-in. frame which, at the appropriate time, could be swung down like a trap door in front of the *I*-figure, and locked into place. All stimuli—except A and B in Fig. 1, which were photographs—were

FIG. 1. A: ILLUSORY SPIRAL; B: ARCHIMEDES SPIRAL; C: INSPECTION-FIGURE WHICH SERVED AS A CONTROL

drawn in black lettering ink on white posterboard, and reflected approximately 20 ft.-L., as measured from *S*'s seated viewing position 6 ft. away. Light was from a 60-w. incandescent bulb in a goose neck lamp placed slightly to the side and about 2 ft. in front of the stimulus-object. A metal shade on *S*'s side of the bulb shielded the light.

The *I*-figure was always spun for 60 sec., as measured by a stopwatch, at which point *E* immediately lowered the *T*-figure.

(a) Condition I. To determine the efficacy of the instructions, the classic Archimedes spiral (B, Fig. 1) was spun at 16 r.p.m. followed by the circle *T*-figure (A, Fig. 3).

Results. All 15 *Ss* reported contraction of the *I*-figure and expansion of the *T*-figure. Common reports of the *I*-figure were: "going in, spinning inward, going far away and it comes back, circles seem to be getting smaller and going toward the center, going toward the middle."[3]

[3] Throughout the experiment, where *Ss* reported lines going toward the center and coming out again, post-test inquiry determined the predominant movement. If *S* replied "mostly in," contraction was scored. If movements were equal, *i.e.* "like an accordian, in and out," no contraction or expansion was scored.

(b) Condition II. It is possible that the expansion after-effect is a result of the sudden introduction of the stationary circles, no matter what the structure of the spinning *I*-figure may be. To test for this possibility, an *I*-figure consisting of circles (C, Fig. 1) was spun at 16 r.p.m., with the *T*-figure as in Condition I *i.e.* A of Fig. 3. The circles of the *I*- and *T*-figures were of exactly the same diameter.

Results. No *S* reported contraction of the *I*-figure, and one *S* reported expansion. Two of the 15 *S*s reported the *T*-figure as expanding, one as

FIG. 2. FRAGMENTED CIRCULAR INSPECTION-FIGURES DERIVED FROM ILLUSORY SPIRAL

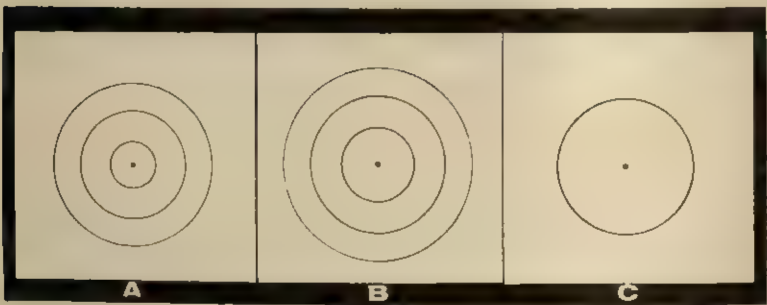

FIG. 3. TEST FIGURES

contracting. This indicates that expansion after-effects are not merely a function of the sudden introduction of the *T*-figure. It also provides a baseline (2 out of 15) for measuring the reliability of results in other conditions.

(c) Condition III. This was the critical condition to determine whether the spinning fragmented circles of the *I*-figure (A, Fig. 2) were perceived as contracting, and whether they produced an expansion after-effect on a *T*-figure composed of stationary concentric circles of the same diameter (A, Fig. 3). Again, the *I*-figure was spun at 16 r.p.m.

Results. No *S* reported contraction of the *I*-figure. All 15 *S*s reported expansion of the *T*-figure. This is clear evidence that after-effects of expansion can be induced without perception of inward movement of the inducing figure.

The spinning fragmented circles were variously described as: "The lines are moving, oscillating; wriggling in and out; like a pinwheel; lines are vibrating back and forth; resonating, wavy lines tend to disappear; lines become a solid circle; like a saw going around; getting smaller, and larger too; just lines going around."

(d) Condition IV. To investigate the effects of speed, a new group was tested as in Condition III, but with the *I*-figure spun at 78 r.p.m.

Results. Two of 15 *S*s reported contraction of the *I*-figure. Eleven of 15 reported expansion of the *T*-figure, a reliably ($P < 0.01$) greater number than had reported expansion after effects in Condition II, as determined by the Fisher-Yates test for 2×2 contingency tables.[4] Two-tailed statistical tests are used throughout.

Although four fewer *S*s reported the after-effect after 78 r.p.m. than after 16 r.p.m. of the *I*-figure, as in Condition III, these differences are not reliable ($P = 0.10$). Thus, there is only suggestive evidence that accelerating the spin of the fragmented circles decreases the perception of expansion after-effects.

(e) Condition V. To determine whether after-effects of contraction can be induced by a fragmented circle, a mirror image of the previously used *I*-figure was drawn (B, Fig. 2). The *T*-figure was the same used in Conditions I–IV, with its circles of the same diameter as those of the *I*-figure. For this and the final two conditions, the *I*-figure was spun at 16 r.p.m.

Results. Three *S*s reported expansion of the *I*-figure. Two *S*s reported contraction and two others reported expansion of the *T*-figure. After-effects of contraction have not been demonstrated.

EXPERIMENT II: FIELD-EFFECTS

Available evidence from experiments with the Archimedes spiral indicates that after-effects of expansion are specific to the area stimulated by the spinning *I*-figure.[5] The structure of the fragmented circles, however, allows for a unique test of whether field-effects occur *between* stimulated areas. Since the spinning fragmented circles do not play continuously

---

[4] D. J. Finney, The Fisher-Yates test of significance in 2 x 2 contingency tables, *Psychometrika*, 35, 1948, 145-156.

[5] B. C. Griffith and H. H. Spitz, Some observations on the spiral after-effect, this JOURNAL, 72, 1959, 139-140.

across the entire disk, as in the Archimedes spiral, a *T*-figure can be constructed with contours falling between the contours of the inducing *I*-figure.[6]

(f) Condition VI. The same fragmented circles served as an *I*-figure (A, Fig. 2). The *T*-figure consisted of three concentric circles, each one slightly larger than the three fragmented circles of the *I*-figure, so that the inner two circles fell midway between areas of stimulation (see B, Fig. 3).

Results. Four of the 15 *S*s reported contraction of the *I*-figure. Eight *S*s (two of whom had reported contraction of the *I*-figure) reported after-effects of expansion. When compared with control Condition II, this ratio of reported after-effects is of marginal reliability ($0.05 > P < 0.10$). There are significantly fewer ($P = 0.01$) field after-effects than direct after-effects, as in Condition III.

These results suggest the possibility that after-effects of field-movement may indeed exist, though in lesser magnitude than direct after-effects. The final two conditions will throw additional light on this problem.

(g) Condition VII. This condition was introduced to determine whether a *T*-figure made up of a single circle would exhibit after-effects of field-movement. The *T*-figure consisted of a circle (C, Fig. 3) with a diameter midway between the outer two fragmented circles of the *I*-figure (A, Fig. 2).

Results. Three of 15 *S*s reported that the *I*-figure contracted. Eleven *S*s (two of whom had reported contraction of the *I*-figure) reported expansion of the *T*-figure, thereby confirming the presence of after-effects of field-movement, suggested by the results of Condition VI.

(h) Condition VIII. The final condition was to determine whether field-effects could be induced *outside* the stimulated area. For this purpose, a new fragmented circle figure was constructed, identical to the disk used previously (A, Fig. 2), but with the outer circle eliminated (C, Fig. 2). The *T*-figure was the same one used in Condition VII (C, Fig. 3).

Results. One of 15 *S*s reported contraction of the *I*-figure. Three of the 15 *S*s reported expansion of the *T*-figure, indicating that this after-effect was not demonstrated. Thus, there are siginficantly fewer ($P = 0.01$) reports of expansion when the contour of the *T*-circle lies outside than when it lies between the contours of the inducing fragmented circles. These results suggest that field-effects of certain after-effects of movement may occur between, but not outside, stimulated areas.

---

[6] An Archimedes spiral could, in fact, be adapted to this arrangement by merely placing a thick, stationary blank circle of medium diameter over the spinning spiral, and using as a *T*-figure a black outline circle whose contours would fall within the contours of the stationary circle.

Discussion

If all *S*s from Conditions III, IV, and VI are combined, only 7 of 45 reported contraction of the spinning fragmented circles, while 34 of 45 reported expansive after-effects of the stationary, true circles. This finding, that expansion after-effects can be induced without the phenomenal perception of contraction of the inducing figure, is a novel one. The most plausible explanation would seem to be that—when one fixates the center of the spinning fragmented circles—as each tilted line crosses any single point on its circumference, it is moving in a slight inward direction. This slight inward movement is below the threshold of phenomenal recognition, but the repeated passage of the lines induces cumulative neural effects which act upon the stationary *T*-figure. As the speed of spin is increased, the lines tend to blend, analogous to flicker fusion, and the full effect of each separate tilted line is not as readily registered.

The same amount of *outward* tilt is, however, below threshold for inducing contractive after-effects. This is congruent with previous findings using the Archimedes spiral, where after-effects of expansion are more readily induced then after-effects of contraction.[7] It has been suggested that the spiral used to induce after-effects of contraction moves toward the rim of the disk, dissipating certain neural effects outside the area of immediate stimulation. With a spiral which produces an after-effect of expansion, no such dissipation can occur, and thus its effect is greater.[8] In view of the present results, this explanation no longer seems tenable.

There is, however, another possibility. Almost from birth, we are frequently moving or being carried forward, and then stopping. Therefore, the world about us is frequently expanding, and becoming stationary, much like an expanding *I*-figure becoming its own stationary *T*-figure. It is possible that small after-effects of contraction continually occurred, and that we have gradually developed some adaptation for these effects, just as short-term, repeated after-effects of movement have shown some effect of adaptation over short periods of time.[9] On the other hand, although objects frequently recede from us, there has been much less opportunity for contraction of the entire field in a symmetrical manner. For this reason,

---

[7] For an excellent review of this problem, see Paul Bakan and Kiyoe Mizusawa, Effect of inspection time and direction of rotation on a generalized form of the spiral after effect, *J. exp. Psychol.*, 65, 1963, 583-586.

[8] C. G. Costello, Massed practice and spiral after-effect. *Percept. mot. Skills*, 12, 1961, 11-14. H. H. Spitz, Neural satiation in the spiral after-effect and similar movement aftereffects. *ibid.*, 8, 1958, 207-213.

[9] Costello, *op. cit.*, 12; H. J. Eysenck and S. B. G. Eysenck, Reminiscence on the spiral after-effect as a function of length of rest and number of pre-rest trials, *Percept. mot. Skills*, 10, 1960, 93-94.

after-effects of expansion may be less adapted, and consequently occur at lower thresholds. If this theory is correct, then after-effects of expansion and contraction should not differ when the fixation-point is peripheral to the spiral. Preliminary findings indicate that this indeed is the case.

The presence of field-effects is a unique finding. If further confirmed, it must be taken into account in any attempt to give a theoretical physiological explanation of after-effects of movement.

Summary

A new figure consisting of fragmented circles was abstracted from an illusory spiral. These fragmented circles did not create the impression of contraction while spinning, but did produce an after-effect expansion on an appropriate stationary test-figure. After-effects of contraction, however, could not be induced. Possible explanations for these results were posited. Field effects on after-effects of movement were also reported.

VARIABLES AFFECTING REVERSAL-SHIFTS IN YOUNG CHILDREN

By W. E. JEFFREY, University of California, Los Angeles

To account for the fact that animals and young children find reversal-shifts more difficult while older children and adults find reversal-shifts easier than nonreversal-shifts,[1] Kendler and Kendler propose a perceptual or verbal mediational mechanism.[2] Although, as they point out, the validity of the mediational mechanism does not depend on its being coördinated with observable events, one is tempted to try to do so.[3] Kendler and Kendler have noted, for example, a correlation between language-usage and the learning of reversal-shifts,[4] and have tried to facilitate reversal-shifts by requiring children to label the stimulus-components to which they have learned to respond.[5] If reversal-shifting was the result of the availability of appropriate labels, however, it would be hard to account for the fact that in the experiment by Kendler, Kendler, and Learnard, even the 10-yr. old *Ss* made fewer than 65% of reversal-shifts.[6] Furthermore, the theory predicts that nonreversal-shifts would decrease with age as the result of increased mediation, but in fact it was the inconsistent responding that decreased as reversal-shifting increased.[7] With these inconsistencies in mind, it would seem reasonable to look to other aspects of the situation for an explanation of the developmental relationship.

---

* Received for publication February 4, 1965. This research was supported in part by grants from the National Science Foundation and the University of California Committee on Research. The author is indebted to Mr. Barry Godolphin for his able assistance, and to Mrs. Ruth Brody and Mrs. Margaret Mathews for their coöperation in making their pupils available as subjects.

[1] R. T. Kelleher, Discrimination learning as a function of reversal and nonreversal shifts, *J. exp. Psychol.*, 51, 1956, 379-384; T. S. Kendler and H. H. Kendler, Reversal and nonreversal shifts in kindergarten children, *ibid.*, 58, 1959, 56-60; T. S. Kendler, H. H. Kendler, and Beulah Learnard, Mediated responses to size and brightness as a function of age, this JOURNAL, 75, 1962, 571-586.

[2] Kendler and Kendler, Vertical and horizontal processes in problem solving, *Psychol. Rev.*, 69, 1962, 1-16.

[3] Kendler and Kendler, *op. cit.*, 7.

[4] Kendler, Kendler, and Learnard, *op. cit.*, this JOURNAL, 582.

[5] T. S. Kendler, H. H. Kendler, and Doris Wells, Reversal and nonreversal shifts in nursery school children, *J. comp. physiol. Psychol.*, 53, 1960, 83-87.

[6] Kendler, Kendler, and Learnard, *op. cit.*, this JOURNAL.

[7] Kendler and Kendler, *op. cit.*, *Psychol. Rev.*, 69, 1962, 1-16.

Experiment I

Learning-rate is one variable that may be considered. Kendler, Kendler, and Learnard report a positive relationship between rate of learning in the first phase of their discrimination-task and the likelihood of a choice indicating a reversal-shift in the testing phase.[8] It is a common observation that children often learn simple discriminations slowly because of their responses to irrelevant aspects of the learning-situation. Although one function of verbal labels might be that they limit responses to irrelevant cues, Kendler, Kendler, and Wells did not find that labels facilitated reversal-shifts when *S* was asked to use them on the last 10 trials prior to having to learn either a reversal- or nonreversal-shift.[9] It is not clear why that treatment was not successful. It would be equally interesting, however, to equate groups for learning-rate by reducing responses to irrelevant cues without training in labeling. The following experiment was done with learning-set training to accomplish this purpose.

Method. The method used duplicates that of Kendler, Kendler, and Learnard. It involved two training phases and one testing phase. In Phase I, two pairs of stimuli, either a large white and small black square, or a large black and small white square, were presented sequentially, and *S* was reinforced for selecting a square in each pair on the basis of one of four stimulus-components, *e.g.* black, white, large, or small. After *S* had met a criterion in Phase I, one of the above pairs was no longer presented, and *S* was reinforced for choosing the previously negative member of the remaining pair. For example, if 'large' had been the positive component in Phase I, *S* might be reinforced for selecting the small black square rather than the large white one in Phase II. Inasmuch as only one pair of squares was present in Phase II, the training was ambiguous in that the solution could be accomplished on the basis of either of the small or the black component. The component *S* utilized presumably is ascertained after *S* has met a criterion in Phase II by presenting the second pair of stimuli (large black and small white). If *S* chooses the small white stimulus, it is assumed that he has been responding in terms of size in Phase II, and therefore he is said to have made a reversal-shift. If he chooses large black, it is assumed that he has been responding in terms of brightness in Phase II and he is said to have made a nonreversal-shift.

Subjects. The *S*s were 42 children with a mean age of 10.5 yr. attending the University Elementary School, and 16 children with a mean age of 4.25 yr. attending the Children's World Nursery School in Los Angeles. The children in each age-group were assigned at random to the learning-set or control conditions prior to serving in the experiment.

Apparatus. A modified Wisconsin General Test Apparatus was used to present the stimuli. The holes for marbles used as reinforcers were 7½ in. apart. The stimuli were black and white circles 2 1/16 and 3⅞ in. in diameter and ½ in. high. There were two complete sets which were alternated from time to time during

---

[8] Kendler, Kendler, and Learnard, *op. cit.,* this Journal, 580.

[9] Kendler, Kendler, and Wells, *op. cit.,* 83-87.

training. There were also 20 pairs of stereometric forms for learning-set training which were made up of small toys affixed to masonite squares which covered the reinforcement-holes. *E* was completely shielded from *S*, but could observe *S*'s response through a slot in the panel between them.

Procedure. The *Ss* were tested individually either in a laboratory room adjacent to the University Elementary School or in a trailer-laboratory adjacent to the play-yard of the nursery school. The instructions indicated that a marble would be hidden under one of the two blocks or toys, and that *S* was to try to discover where the marble was hidden to obtain as many marbles as possible. When there was no marble under the "toy" or "block," *S* was to return a marble by placing it under the "toy" or "block" he had chosen. The marbles were collected in a cup, and at the termination of the experiment the nursery school children exchanged their marbles for a small toy which they had chosen at the beginning of training.

For those *Ss* given learning-set (*LS*) training, pairs of stimuli were presented for six trials in an object-quality discrimination. The correct objects and the sides on which the correct objects were placed were counterbalanced on the first trial and random on subsequent trials. *Ss* were trained until they reached a criterion of 9 correct second-trial responses in 10 sets. At this point, Phase I of the Kendler-Kendler-Learnard procedure was begun. A control group (*C*) had no training prior to Phase I.

The circular black and white blocks served as discriminanda and a counterbalanced design reinforced responses to either the large, small, black, or white components. It should be noted that the stimulus-set included four blocks which were presented in such a way that each pair always differed in both dimensions. The two pairs of stimuli were presented equally often and never more than twice in succession. The correct stimulus was presented equally often on each side and never more than twice in succession on one side.

After *S* reached the criterion of 9/10 correct choices in Phase I, Phase II began without interruption. *S* now was presented with only a single stimulus-pair. If he had learned to respond to 'large' regardless of brightness in Phase I, he was reinforced in Phase II for responding to 'small.' The criterion in Phase II, could, however, be met on the basis of either size or color (*i.e.* small or white). After reaching a criterion of 9/10 correct choices, the basis for *S*'s response in Phase II was assessed in Phase III when the second set of blocks was randomly interspersed for 10 presentations in the next 18 trials. All responses were reinforced in Phase III.

Results. For the 10-yr.-old control *Ss*, the mean number of trials to criterion was 24.3 in Phase I and 6.7 in Phase II; for the learning-set *Ss*, the respective means were 4 and 3. The difference for Phase I was significant ($t = 3.11$, $p < 0.001$); for Phase II, it was not. The 4-yr.-old *Ss* took 28 and 8 trials without *LS*-training and 2 and 4.2 trials with *LS*-training. Here again the difference was significant for Phase 1 ($t = 3.75$, $p < 0.001$), but not for Phase II.

Of main concern is the performance of the *Ss* in Phase III. Children who made eight or more responses to one test-stimulus in the 10 presentations were classified as reversal (*R*) or nonreversal (*NR*) *Ss* according to their choice, while *Ss* who made fewer than eight responses to a

single stimulus were classified as inconsistent (I). If the equalization of learning-rates effectively alter reversal-shift performance, it might be expected to show up most clearly in a comparison of the 4-yr.-old *LS*- and *C*-groups in that the 4-yr.-old *LS*-group should perform more like the 10-yr.-old *C*-group and thus should differ significantly from the 4-yr-old *C*-group. The difference between the 10-yr.-old groups would leave less room for change.

As may be seen in Table I, there was no difference in reversal-shifts between the two younger groups. Furthermore, the difference between the two older groups was opposite to expectation. The low frequency of *I*-responses made it necessary to combine *I*- and *NR*-categories for a Chi-square test of the differences in frequency of reversal-shifts. Only two differences needed testing. First, these was the difference between the older groups, which was not significant. Secondly, there was the question of the developmental effect. When the treatment-groups within each age-level were combined, a chi-square for age of 3.10 ($p < 0.10$) was obtained. Thus, although the *LS*-training clearly equated the rate of learning for the two age-groups, it did not increase the proportion of reversal-shifts either in the younger or older groups, and the age-difference in reversal-shifting remained.

TABLE I

EFFECT OF LEARNING-SET (*LS*) OR NO PRETRAINING (*C*) ON 4- AND 10-YR.-OLD CHILDREN'S PERFORMANCE ON THE TEST-TASK AS MEASURED BY THE PERCENTAGE OF RESPONSES IN EACH CATEGORY

	4-yr-olds		10-yr-olds	
	LS (N=8)	*C* (N=8)	*LS* (N=21)	*C* (N=21)
Reversal-shift (*R*)	37.5	37.5	66.7	81.0
Nonreversal-shift (*NR*)	25.0	50.0	33.3	9.5
Inconsistent (*I*)	37.5	12.5	0.0	9.5

EXPERIMENT II

While observing the essentially errorless performance of the *LS*-groups, *E* had the idea that a number of the younger *S*s probably were making nonreversal-shifts, not because of a failure to mediate, but because they were using a different mediator than the older *S*s. They appeared to respond to both components of the reinforced stimuli. To illustrate, an *S* might identify the two correct stimuli in Phase I as (1) large black, and (2) large white, without ever reducing the identification to the single common characteristics (large). In Phase II, *S* learned that the large black stimulus was no longer correct and that the small white one was now

reinforced. Because *S* had not previously used 'large' as the single appropriate label, he would probably not use 'small' in such a capacity after Phase-II training. Therefore, when the other pair of stimuli from Phase I was reintroduced, the *S* would continue to respond to it as he had learned to do originally, making what would be identified as a nonreversal shift.

On the assumption that some *S*s were responding in this way, it was proposed that the likelihood that the *S*s would respond to the test-stimuli as the identical Phase-I forms would be reduced if the form of the stimuli was changed (from circles to squares) after Phase-I training. The consequent reduction in generalization should also reduce the proportion of nonreversal-shifts.

Subjects. A total of 85 *S*s served in this study. There were 25 *S*s averaging 4.16 yr. (range = 46–54 mo.) from the Children's World Nursery School in Los

TABLE II

PERFORMANCE OF DIFFERENT AGE-GROUPS ON THE TEST-TASK (AS MEASURED BY THE PERCENTAGE OF RESPONSES IN EACH CATEGORY) WHEN FORM OF STIMULI WAS SHIFTED FROM PHASE I TO II IN EXPERIMENT II

	4-yr-olds ($N=25$)	6-yr.-olds ($N=20$)	8-yr.-olds ($N=20$)	College students ($N=20$)
Reversal-shift (*R*)	76	75	80	75
Nonreversal-shift (*NR*)	16	15	15	15
Inconsistent (*I*)	8	10	5	10

Angeles. There were 20 6-yr.-old and 20 8-yr.-old children from the University Elementary School, and 20 college students from an introductory psychology class in the University.

Apparatus. The apparatus and stimuli were the same as in Experiment I, except for the addition of two sets of square stimuli which duplicated the circular stimuli in area and brightness. Half of each group started with square blocks and shifted to round blocks in Phase II; the other half started with round and shifted to square blocks.

Procedure. The instructions were the same as in Experiment I. The 4-yr.-olds received a toy at the completion of the experiment. The other *S*s worked only for the marbles, which were returned to *E* at the conclusion of the session.

Results. *S*s who made eight or more responses to one test-stimulus in 10 test-presentations were classified as a reversal (*R*) or nonreversal (*NR*) according to their choices. *S*s who made fewer than eight responses to a single stimulus were classified as inconsistent (*I*).

The proportion of *S*s in each group falling in each category is presented in Table II. All proportions for each type of shift are remarkably constant across the various age-levels. Thus, the change in the form of the stimuli

does modify the behavior of the younger *S*s in such a way as to destroy the developmental function.

Comparing the 4-yr.-old *S*s of this experiment with those of Experiment I (see Table I), we note that the change in form during Phase II reduced both the nonreversal shifts and the inconsistent behavior. If the *NR*- and *I*-categories are pooled to increase the cell-frequencies to an acceptable levl, a Chi-square of 6.05 ($p < 0.02$) is obtained, indicating that the category-proportions are not independent of treatment. If this effect is due to the reduction of a tendency to respond to the test-pair in Phase III on the basis of similarity to the stimulus-pair of Phase I, then the normally poorer reversal-shift performance of the younger *S*s is apparently due to their greater tendency to respond to the specific stimulus-compounds, rather than to the common component. In other words, the younger *S* make nonreversal-shifts, not because he lacks mediators, but because he does not abstract the single cue or use the mediator that the experimenter has chosen to reinforce. This interpretation suggests a reason for the failure of the *LS*-training to modify the behavior of the younger *S*s in Experiment I. Although *LS*-training eliminates some irrelevant hypotheses, such as position or sequence, and forces attention to the reinforced objects, it does not require attention to a specific stimulus-components.

The manipulation of verbal mediators may, however, bring about attention to specific components. For example, Kendler has demonstrated that the percentage of reversal-shifts is increased if kindergarten children are required to verbalize labels for the positive and negative stimuli prior to making a choice.[10] Although Kendler does not specify very precisely what the effect of the training in labeling is, Experiment II suggests that its main role may be to inhibit responses to irrelevant dimensions. That the labels exercise relatively tenuous control as more specific mediators is indicated by the fact that some of Kendler's children continued to report "black is the winner and white is the loser" even after successfully shifting their choices to 'white' in Phase II. It would have been interesting if Kendler had studied a group which learned to say "large white is the winner and small black is the loser." According to the hypothesis presented here, children so trained would have shown a much lower percentage of reversal-shifts in that such training would maintain or possibly even strengthen the tendency which is presumed here to account for many of the nonreversal-shifts.

---

[10] T. S. Kendler, Verbalization and optional reversal shifts among kindergarten children, *J. verb. Learn, verb. Behav.*, 3, 1964, 428-435.

Although the present hypothesis provides a plausible explanation for *S*'s choice of the nonreversal-shift stimulus, it still is necessary to explain why, once *S* no longer identifies the test-pair with the similar Phase I-set, he now chooses the reversal-shift stimulus. In other words, a two-component *S* would probably still respond "small white" in Phase II, and, even when a change in form reduces the tendency to choose large white, the tendency to choose small black is not increased. If one assumes, however, that only 50% of the *S*s respond to two components (Experiment I), and that, when their tendency to generalize to the Phase-I pair is inhibited, their choice of test-stimulus is a matter of chance, then one would expect 25% of this group to make reversal-shifts. If one also assumes that the single-component *S*s all make reversal-shifts, then the total reversal-performance should be 75%, which is the percentage obtained. When Kendler forced verbal labels to single components, the responses to the irrelevant dimensions were sufficiently inhibited that a nonreversal-shift in the test-trials were even less likely, and she obtained from 85-100% reversal-shift performance.

In conclusion, although the testing procedure used in these experiments is appealing in that it appears to indicate the basis for each *S*'s response, it is obvious that a number of variables may influence children's choices in such a situation. It also seems clear that, in studying the behavior of children of 4 yr. or older, it probably is more important to ask what role mediators play than whether they are used.

Summary

Two experiments were done to investigate alternatives to a verbal-mediation interpretation of the developmental function obtained for reversal-shift behavior. In Experiment I, half of a group of 4-yr.-olds and half a group of 10-yr.-olds were given learning-set training prior to discrimination. This training equated learning-rate in the two age-groups, but it did not affect the developmental function (the 4-yr.-olds continued to show a lower proportion of reversal-shifts than the 10-yr.-olds.) In Experiment II, the form of the stimuli was changed from the first to the second phase of training, which did destroy the developmental function. This result suggests that the developmental function reflects the younger childrens' failure to abstract a single dimension rather than a lack of verbal mediation.

THE RELATIONSHIP BETWEEN CRITICAL FLICKER FREQUENCY (*CFF*) AND SEVERAL INTELLECTUAL MEASURES

By Rudolph L. Zlody
Holy Cross College

During the last few decades, there has been disagreement over the relative roles of peripheral versus central factors in the resolution of intermittent photic stimulation. The peripheral view stresses inertia of the retinal mechanism; those who hold to central determination of critical flicker frequency (*CFF*) place it, as Teuber did, in the occipital lobe.[1] Halstead, who also holds to cortical resolution, places the site in the frontal lobes and states that *CFF* is related to cerebral metabolism and thus reflects the functional status of the central nervous system.[2]

In his efforts to make a contribution to "quantitative biology," Halstead attempted to relate changes in *CFF* to circumscribed injuries to various portions of the brain. He felt that by contrasting test-results of *S*s with an intact and uninjured central nervous system (*CNS*) with those whose histories indicated cortical damage, the principles of localization of function could be placed on a more objective basis.

Against test-results of a large group of *S*s who had undergone leucotomy, Halstead compared the data from normal, neurotic, "miscellaneous" controls, and 50 healthy adult men who had recently recovered from a concussive type of head injury. A factor analysis of these data produced four factors of "biological intelligence." The *P* or 'Power factor' was heavily represented by *CFF*, and was interpreted as a dynamic factor which probably best reflects the over-all status of the brain.

Since *CFF* was lower in the brain-injured, Halstead argued that the retina could be eliminated as a factor in *CFF*. Furthermore, he stated, biological intelligence and *CFF* are maximally mediated in the frontal lobes. In a second study he found that all four factors correlate in the 0.30s with both verbal and performance psychometric intelligence.[3] On this basis, he suggested a positive relationship between *CFF*, a component of biological intelligence, and psychometric intelligence.

---

* Received for prior publication June 10, 1965. Submitted to the Graduate Faculty of Fordham University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

[1] Hans-Lukas Teuber and M. B. Bender. Changes in visual perception of flicker, apparent motion, and real motion after cerebral lesion, *Amer. Psychol.*, 3, 1948, 246-247.

[2] W. C. Halstead, *Brain and intelligence*, 1947.

[3] Halstead, Biological intelligence, *J. Pers.*, 20, 1951, 118-130.

Young, in her analysis of the 72 tests used in the Columbia-Greystone I Project, found *CFF* to correlate with none of the other tests in a meaningful manner, nor could she corroborate the existence of a *P* factor.[4] Further, *CFF* was not affected significantly by surgery on the frontal cortex.

Tanner found a moderately high correlation between *CFF* and the American Council on Education intelligence test (ACE) among a group of college students.[5] In a second study in which he lowered the intensity of the light stimulus, he found 5 of 12 correlations significant at the 0.05 level.

Proceeding on a wealth of evidence that older persons have lower *CFF* than younger people, Colgan felt that *CFF* may prove to be a valuable tool in measuring the process of aging.[6] In his study of 40 men between ages 65 and 95, in which he found marked decreases with age, he felt, as Tanner had suggested earlier, that his groups differed in more than age alone. Colgan subsequently tested his *S*s on the Wechsler-Bellevue (W-B), Form I. Five of the 14 correlations between *CFF* and *W-B* were significant at the 0.01 level, and 6 more at the 0.05 level. By holding age constant, he found that the correlation of *CFF* with the *W-B* was significant at 0.05; the correlation between age and *CFF*, with intelligence partialled out, was not significant.

Colgan's second investigation involved *CFF* and the American Council on Education intelligence test (*ACE*) scores of 40 men (college students).[7] None of these correlations reached significance. He did not consider his findings to be a contradiction of Tanner because differences in *CFF* of his two groups may be due to intelligence rather than age. Since a significant relationship exists between *CFF* and intelligence independent of age, all studies which have included age and *CFF* as first-order variables "must be reëxamined to determine whether intelligence was a controlled factor—if it was not, the conclusions of those studies must be qualified accordingly."[8]

Analysis of a number of studies by Landis and his associates revealed no relationship of *CFF* to intelligence, and Landis and Hamwi suggest that "the correlations of Tanner and Colgan are examples of random fluctuation."[9]

There appears to be some evidence to support the contention that a positive relationship exists between *CFF* and intelligence in the aged. Loranger concluded that although intelligence as we know it depends on factors other than brain function, dependence of intelligence on cortical function is clearly illustrated in the aged.[10] Since evidence for decreased neural efficiency does not exist before the age of physical decline, *CFF* in the aged may be, as Loranger states, a neurophysiological correlate of the higher intellectual functions.

---

[4] K. M. Young, Critical flicker frequency, in F. A. Mettler (ed.), *Selective Partial Ablation of the Frontal Cortex*, 1949.

[5] W. P. Tanner, Jr., A preliminary investigation of the relationship between visual fusion of intermittent light and intelligence, *Science*, 112, 1950, 201-203.

[6] C. M. Colgan, Critical flicker frequency, age, and intelligence, this JOURNAL, 67, 1954, 711-714.

[7] Colgan, The relationship between critical flicker frequency and several psychological variables, Unpublished Doctoral dissertation, University of Florida, 1954.

[8] Colgan, *op. cit.*, 713.

[9] Carney Landis and Violet Hamwi, Critical flicker frequency, age, and intelligence, this JOURNAL, 69, 1956, 459-461.

[10] A. W. Loranger, Critical flicker frequency and some intellectual functions in old age, Unpublished Ph.D. dissertation, Fordham University, 1958.

Using various tests of intellectual ability which she felt were dependent on neural efficiency, Cross was unable to find a significant correlation between *CFF* and intelligence among children ages 6 through 12 yr.[11]

Wilson, using an episcotister, tested 96 *Ss*.[12] His findings suggest that both age and nonverbal intelligence are determinants of *CFF:* nonverbal intelligence and age correlate +0.62, and nonverbal intelligence and *CFF* correlate +0.048.

Since most of the earlier studies have certain limitations in procedure, method, instrumentation, and personal characteristics of the *Ss*, the functional relationship between *CFF* and intelligence must still be regarded as tentative. It was because of this that the study reported here was undertaken.

Method and Procedure

Apparatus. The apparatus used in the present study was an electronic pulse generator designed and built by Ireland.[13] A Sylvania glow modulator tube, type R 59/1130 B, provided a continuous spectrum white light of 2.5 ft.-c. measured at the artificial pupil. A milk glass filter, 1 in. from the lens, diffused the light. When the test-patch was viewed through an artificial pupil of 1.6 mm. in diameter from a distance of 14.33 in., the visual angle of 1°20′ subtended in the retina ensured foveal stimulation. Viewing was with the dominant eye.

The apparatus provided a linear calibration with a frequency range from 4–60 c.p.s. Light-dark ratio (*LDR*), intensity, and frequency could be controlled and varied independently. Low intensity was produced by use of a one-log Kodak-Wratten neutral-density filter. An *LDR* of 0.5 was used throughout.

Subjects. One hundred healthy men, college students between the ages of 18 and 21 yr. were tested on *CFF*, the *ACE*, and the Intermediate form of the *Primary Mental Abilities* (*PMA*). The mean age of the *Ss* was 20.0 yr. Care was taken in the selection of *Ss* to accept only those who were free of present illness, who had no diseases or disturbances that affected the cardiovascular or neurocirculatory systems, and who had no pathologies or anomalies of the visual mechanism. Finally, screening ruled out those who had had concussive-type head injuries that resulted in unconsciousness, or any pathology of the *CNS*. The *Ss* were not tested for *CFF* when under psychological or physiological stress.

The age-range was deliberately limited because, although there is evidence that *CFF* declines with age, there is also agreement that it shows no decline between ages 17 and 21 yrs.

Procedure. Two levels of intensity were used: 2.5 ft.-c., and an intensity lower by one log unit. The method of limits was used exclusively. The determinations were made at each intensity-level, five ascending and five descending series. The *Ss* were tested on both intensity-levels on two successive days at the same time each

---

[11] Jane P. Cross, Relation of age and mental growth to the *CFF* response in children, *Child Develop.*, 34, 1963, 739-744.

[12] T. R. Wilson, Flicker fusion frequency, age, and intelligence, *Gerontologia*, 7, 1963, 200-208.

[13] F. H. Ireland, Pulsed light measures flicker perception, *Electronics*, 26, 1953, 90-93.

day. To eliminate a practice effect, the first *S* was tested high-low (*H-L*) on the first day, and *L-H* on the second day; the second *S* was tested *L-H* and then *H-L*. No statistical difference was found between ascending and descending thresholds within individuals and both readings were combined into a single *CFF* threshold. It was also determined statistically that *CFF* does not change over 24 hr.

The psychological tests were so chosen as to provide measures which had psychological meaning in themselves, were measures of 'pure' mental factors, represented a variety of tasks with only slight overlap, and were applicable to the age-range and educational level of the *Ss*. The *ACE* was selected because of its high reliability and because present findings could be compared with those of Tanner. The Primary Mental Abilities test, (*PMA*) Intermediate form, was chosen because of the factorial purity of its measures.

The scores on Verbal Comprehension of the *PMA* were skewed so a square root transformation was performed. The means in the present study are higher than the means in the respective manuals, but the *SDs* are only slightly larger. In all but two of the eight subtests the present scores spanned almost the entire range of

TABLE I

MEANS, *SDs* AND RANGES OF *CFF* IN CYCLES PER SECOND
($N=100$)

	Mean		*SD*		Range
	This study	Colgan's	This study	Colgan's	This study
CFF-H	38.8	36.9	2.91	1.8	14.4
CFF-L	33.4	—	2.51	—	11.2
AD-H	0.65	—	0.33	—	1.66
AD-L	0.60	—	0.31	—	1.53

scores reported in the manuals, thus demonstrating the heterogeneous makeup of the sample. Moreover, these same six tests plotted in an almost perfect curve.

Altogether 12 measures were obtained from each *S*: 3 on the *ACE*, 5 on the *PMA*, and measures of central tendency and dispersion for each of 2 intensity levels on *CFF*.

RESULTS

The reliabilities of *CFF* over 24 hr. in the present study were 0.972 for *CFF*—high intensity (*CFF-H*) and 0.941 for low intensity (*CFF-L*)

When the present means are compared with means obtained with comparable equipment, the differences are negligible, as are the *SDs*. The Fordham investigators and the Minnesota group found the range of *SD* to be 2.5 to 3.5 c.p.s. The intraindividual *AD*, with a mean of less than 1 c.p.s., has been reported by Brožek and Keys, by Landis, and by Misiak.[14]

[14] Josef Brožek and Ancel Keys, Changes in *FFF* with age, *J. consult. Psychol.*, 9, 1945, 89-90; Landis, Determination of the critical flicker-fusion threshold *Physiol. Rev.*, 34, 1954, 259-286; Henryk Misiak, Age and sex difference in critical flicker frequency, *J. exp. Psychol.*, 37, 1947, 318-332; F. J. McCabe, The effect of light-dark ratio, age, and sex upon critical flicker frequency. Unpublished Doctoral dissertation. Fordham University, 1954.

It will be noted that both *SD* and the range tend to increase with increase in intensity. The correlation of *CFF-H* and *CFF-L* is 0.682, and the corresponding *AD*s correlate 0.394. There are no comparable reports in the literature to which these correlations can be related.

Table II below reports scores on the psychometric tests administered. Colgan's results are included for comparison.

The means for the present study are higher than the normal data, but the *SD*s

TABLE II

MEANS, *SD*s, AND RANGES OF PSYCHOMETRIC TEST-SCORES
($N=100$)

Tests	Mean			*SD*			Range
	Manual	This study	Colgan's	Manual	This study	Colgan's	This study
ACE							
Q	37	46.4	44.0	11.8	12.6	8.2	62
L	59	75.3	71.7	15.9	18.3	12.1	82
I	96	121.7	115.7	24.9	26.9	17.0	110
PMA							
V	27	42.7	—	13	6.8	—	31
S	28	29.4	—	10	10.9	—	55
R	16	21.0	—	6	7.3	—	22
N	20	29.4	—	9	10.2	—	55
W	45	52.1	—	10	12.6	—	64

TABLE III

PRODUCT-MOMENT CORRELATION OF MEAN-*CFF* AND *AD-CFF* WITH EIGHT TESTS OF MENTAL ABILITIES
($N=100$)

	ACE		*PMA*				
	Q	*L*	*V*	*S*	*R*	*N*	*W*
CFF-H	.005	.178	.052	.002	.099	.020	.303*
CFF-L	.102	.228*	.265†	.058	.173	.027	.148
AD-H	.140	.106	.119	.005	.003	.246*	.103
AD-L	.228*	.146	.153	.114	.090	.315†	.099

* Significant at 0.05 level; † Significant at 0.01 level.
For 98 degrees of freedom: P 0.05=0.197; 0.01=0.257.

are close. The range is of particular interest since it spans with two exceptions, almost the entire range of scores in the manuals. With *CFF* readings also having a wide range any relationships that exist between *CFF* and these measures of intelligence will be enhanced.

Product-moment correlations were computed between psychometric scores and means and *AD*s of *CFF*. The total score of the *ACE* was dropped because of its lack of precise meaning. The relationship between *CFF* and mental ability scores is shown in Table III.

The table reveals that *CFF* correlates significantly with *L*, *V*, with *W*, the first two under low illumination. All three are heavily weighted with verbal ability.

Difficulty is encountered in accounting for the significant correlations between

psychometric scores and the *AD*s. Here again low intensity measures are more frequently related to psychometric scores than *CFF-H*. Although the correlations of *AD* with *Q* and *N* are significant and both are measures of numerical ability, the meaning of this relationship is obscure because the *AD*s are relatively unstable measures. Because the *AD* is skewed and may contribute to the high correlation, the *AD*s were transformed by a square root transformation. The significant correlations of *AD-L* and *Q* and *N* remained.

Discussion

Halstead, Tanner, Colgan, and Wilson reported a positive relationship between *CFF* and various measures of mental ability. Colgan, using a second sample that differed from his first, and Landis and his colleagues, have been unable to do so.

Colgan suggests a factor of attentiveness as tying the two together. Tanner says that greater effort and sustained attention are required at low intensity, and this "mental energy" is characteristic of intelligent persons.

The present data suggest that there may be a critical property in light of low intensity that is not evident with a bright light. Again such factors as effort and attention may be the sought-after variables. The high correlation of *CFF-H* and *W* may suggest that psychological perseveration is related to physiological preservation and that if it is difficult to resolve successive visual stimuli it may be equally difficult to shift from a present to a related response. The effects of such personality factors suggested here—perseveration, rigidity, and flexibility—have not been considered in their possible influence on *CFF*.

This explanation does not necessarily hold when the correlation between variability at low-intensity levels and *Q* and *N* are considered. The fact that they deal with numbers is hardly a sufficient interpretation until the mental operations of number ability are more clearly defined.

The only significant correlation that appears to have meaning is *CFF-H* and *W*. Both require rapid association and hence may require faster or more extensive connections in the cortex. Frequency of discharge or cortical neurons, a shorter refractory period, reduced synaptic delay, or facilitation of impulses are all attractive hypotheses. *CFF*, if it is a neural rather than a predominantly photochemical phenomenon in the retina, may also require greater speed of neural transmission or a more diffuse response in the cortex itself. Lack of precise knowledge in this connection makes this explanation tentative and tantalizing only.

One relationship that has not been investigated is the relative role of speed versus power tests and *CFF*. Colgan (1954a) and Loranger (1958) found significant correlations between intelligence scores which were largely power tests and *CFF* in the elderly. Both conclude that this relationship can be ascribed to decreased neural function through normal degenerative processes. Fruitful research opportunities are available here.

Perhaps the significant correlations obtained in the present study have no direct application to central neural processes and are found in such variables as perceptual or personality factors. It is unsafe at this stage to postulate a physiological mechanism to account for the correlations since physiological processes, which function within rather narrow limits and are revealed in their function only after severe impairment, have not as yet been demonstrated as crucial in *CFF*.

The present results may therefore be regarded as equivocal: five of the 28 corre-

lations between *CFF* and selected measures of intellectual ability are significant, but they are so unpatterned that they defy logical and psychological explanation. It must be concluded that the relationship between *CFF* and intelligence test-scores among college-age students is not confirmed, and the results could be, as Landis suggests, "examples of random fluctuation."

Conclusions

Although several investigators report a significant correlation between *CFF* and scores in tests of intelligence, their conclusions are open to question because of differences in the ages, methods, procedures, and personality characteristics of their *Ss*.

One hundred healthy men (college students) were tested on *CFF-H*, *CFF-L*, the *ACE*, and the *PMA*. Comparison of obtained *CFF* results with other investigations revealed a remarkable consistency in mean *CFF*, intraindividual variability, and group variability.

The correlations between *CFF* and scores on tests of intelligence are all positive: of 28 correlations, 5 are statistically significant, 2 of which involve measures of dispersion. Moreover, 3 of these 5 occurred under low illumination, a finding that requires further elucidation.

The present study did not seek to determine the differential correlation of speed- or power-tests since more emphasis was on 'pure' factors of mind. Pursuit of this area may prove valuable.

No meaningful pattern emerged among the significant correlations that were obtained, and there is no proof in the results obtained to confirm the hypothesis that *CFF* and selected measures of psychometric intelligence are related. Tentative explanations have been advanced for the significant correlations that did occur.

DECREMENT OF THE MÜLLER-LYER ILLUSION AS A FUNCTION OF PSYCHOPHYSICAL PROCEDURE

By Nora I. Parker and P. L. Newbigging
McMaster University, Canada

Decrement of the Müller-Lyer (*M-L*) illusion as a function of repeated observations or trials has been observed in a large number of experiments conducted over the past sixty years.[1] It is somewhat surprising therefore, to note in two recently published studies failure to demonstrate the decrement under a variety of experimental conditions.[2] A careful examination of the procedures employed in the experiments reported in these papers would seem to be of interest.

The paper by Eysenck and Slater reports one experiment, while the monograph by Day reports seven. In all experiments, the Brentano-type figure (angles without the joining horizontal line) was employed. In the Eysenck and Slater study, this figure was presented at one end of an enclosed box which provided at the other end an eye-piece or viewing tube through which *S* looked. Day used a similar apparatus for most of his experiments. In different experiments, Day investigated the effect on the decrement of the illusion of high and low figure contrast, 'open' and 'closed' viewing,[3] 'apparent' and 'objective' instructions,[4] large and small figure, fixation and free inspection, eccentric and central fixation,[5] and prolonged fixation. For the groups of *S*s administered these different treatments, the only ones

* Received for publication November 23, 1962. The research for this paper was supported in part by the Defence Research Board of Canada under Grant No. 9401-13. Arlene Vadum and Albert Kuechler assisted in this study.

[1] For example, C. H. Judd, Practice and its effects on the perception of illusion, *Psychol. Rev.*, 9, 1902, 27-39; E. O. Lewis, The effects of practice on perception of the Müller-Lyer illusion, *Brit. J. Psychol.*, 2, 1908, 294-307; H. R. Crosland, H. R. Taylor, and S. J. Newsom, Practice and improvability of the Müller-Lyer illusion in relation to intelligence, *J. gen. Psychol.*, 2, 1929, 290-306; Wolfgang Köhler and Julia Fishback, The destruction of the Müller-Lyer illusion in repeated trials: I. An examination of two theories, *J. exp. Psychol.*, 40, 1950, 267-281; P. T. Mountjoy, Intrasession decrements and intersession recovery to the Müller-Lyer figure, *Percept, mot Skills*, 13, 1961, 41-57.

[2] R. H. Day, The effects of repeated trials and prolonged fixation on error in the Müller-Lyer figure, *Psychol. Monogr.*, 1962, (No. 533), 1-19. H. J. Eysenck and P. Slater, Effects of practice and rest on fluctuations in the Müller-Lyer illusion., *Brit. J. Psychol.*, 49, 1958, 246-256.

[3] 'Closed' viewing involved the use of the box, whereas in 'open' viewing the illusion was mounted on panels set on the end of a table and not enclosed.

[4] In 'apparent' instructions, *S* was to set the variable to apparent equality with the standard. In 'objective' instructions, he was to set it to objective or physical equality.

[5] In eccentric fixation a point above the central apex of the figure was fixated, while in central fixation the central apex itself was fixated.

showing a significant decrement were those instructed to set the variable to physical rather than apparent equality, those who freely inspected a large or small figure, or those who fixated a small figure at its center. A significant decrement also followed prolonged inspection of the *M-L* figure, but the two control groups—one of which fixated a point, while the other kept their eyes closed for a comparable period—also showed a significant decrement.

The experiment by Eysenck and Slater involved only instructions to judge when the variable was apparently equal to the standard, and the interspersing of a fixation- and a rest-period after 40 and 50 trials, respectively.

As noted above, the decrement in the *M-L* figure has been observed in a great many experiments conducted under a wide variety of conditions, some of which come close at least to those employed by Day and by Eysenck and Slater. Contrary to the findings of Day and of Eysenck and Slater, the decrement has been obtained under conditions of closed viewing,[6] when *S*s were instructed to set the variable at apparent equality,[7] and under conditions of fixation.[8]

The one procedural detail, not so far referred to, in which the experiments by Day and the one by Eysenck and Slater differ from most *M-L* experiments reported in the literature is in the psychophysical procedure employed. In Eysenck and Slater's experiment the variable (distance between the "open" obliques) was always set longer than the standard before each trial and was gradually shortened until *S* halted the process by indicating that the two appeared equal. In Day's experiments, the variable (again the distance between the "open" obliques) was always set to physical equality with the standard and adjusted to "apparent" or "objective" equality by either *S* or *E*. This procedure is similar to that of Eysenck and Slater's since the variable was the extent that appears longer at physical equality and one would surmise that *S* typically required it to be shortened, although Day is not explicit on this point.

When the Method of Average Error or Adjustment is used in the study of the *M-L* illusion, it is customary to intermix randomly longer and shorter settings of the variable to control for directional error. It was considered possible that the failure of Day and of Eysenck and Slater to obtain the decrement might be due to the unusual procedures they employed, and for this reason the experiment described below was done.

Method

Subjects. The *S*s were 32 women, volunteers from a local teachers college. They were randomly assigned to one of four groups.

Apparatus. The apparatus for presenting *M-L* figure consisted of a rectangular

---

[6] Mountjoy, *op. cit.*, 50.

[7] Nora I. Parker and P. L. Newbigging, Magnitude and decrement of the Müller-Lyer illusion as a function of pretraining, *Can. J. Psychol.*, 17, 1963, 134-140.

[8] Köhler and Fishback, *op. cit.*, 269.

frame 48 in. x 25 in. which enclosed a plywood panel with a 31 in. x 10 in. opening in its center. The frame and panel were painted white, forming a white surround for the stimulus-figure. Two additional plywood panels slid in grooves at the back of the frame and overlapped slightly. From the front view the two sliding panels completely filled the center opening. The half of the stimulus-figure used as the standard was fixed on the foremost of these panels while the variable half was fixed on the other which slid in behind. Movement of the variable in either direction was accomplished by a reversible electric motor which could be controlled by either *E* or *S*. *S* activated the motor by a remote control key. The motor moved the panel at a constant speed of approximately 7 mm. per sec.

The *M-L* figure was the Brentano-type (angles without the horizontal line) and was drawn in India ink on white cardboard. The width of all lines was 1 mm. The apexes of the angles of the standard, which was the side with the angles turned in, were 16 cm. apart.

Procedure. For all conditions *S* was seated at a table 6.0 ft. in front of the apparatus. He was instructed to move the variable portion of the figure, by depressing the key adjacent to his preferred hand, until it *appeared* equal to the standard. The variable was on the left for half the *Ss* in each group and on the right for the other half. While free viewing of the stimulus-figure was permitted, a chin-rest limited head-movements. Every *S* made a total of 96 seetings in a single session lasting from 30 to 45 min. under one of four procedural conditions.

The procedural conditions, and hence the group of *Ss*, differed in the setting of the variable prior to the *S*'s adjustment. These were: (1) a mixed procedure, where four different longer and four different shorter settings were randomly intermixed for successive trials. Each of the eight settings was such that the variable, prior to adjustment, was either obviously shorter or longer than the standard. (2) In the second procedure, only the four longer settings employed in the mixed procedure were used. This is essentially Eysenck and Slater's procedure.[9] (3) In the third procedure, only the four shorter settings employed in the mixed procedure were used. Finally, (4) the variable was always set to physical equality with the standard before *S*'s adjustment. This last is the procedure used by Day.[10]

RESULTS

Since the interest in this experiment is in the decrement of the *M-L* illusion over trials, the data were treated as follows. For each *S*, the amount of illusion on the second and subsequent block of trials was computed as a percentage of the illusion on the first block. These percentages were then averaged over *Ss* within each procedural condition and are presented graphically in Fig. 1.

An analysis of variance was performed on these percentages, excluding the first block of trials, since the method of analysis had reduced the variance of this block to zero. This analysis is summarized in Table I. The main interest is, of course, in the significant effect of trials and the signifi-

[9] Eysenck and Slater, *op. cit.*, 250.
[10] Day, *op. cit.*, 1.

cant trials × procedures interaction, both of which are apparent in Fig. 1.

To determine which of the four procedures gave a significant decrement over trials, a simple analysis of variance was performed on the data for each procedure. The data used were the error measurements in centimeters for each block of trials, including the first trial. A significant

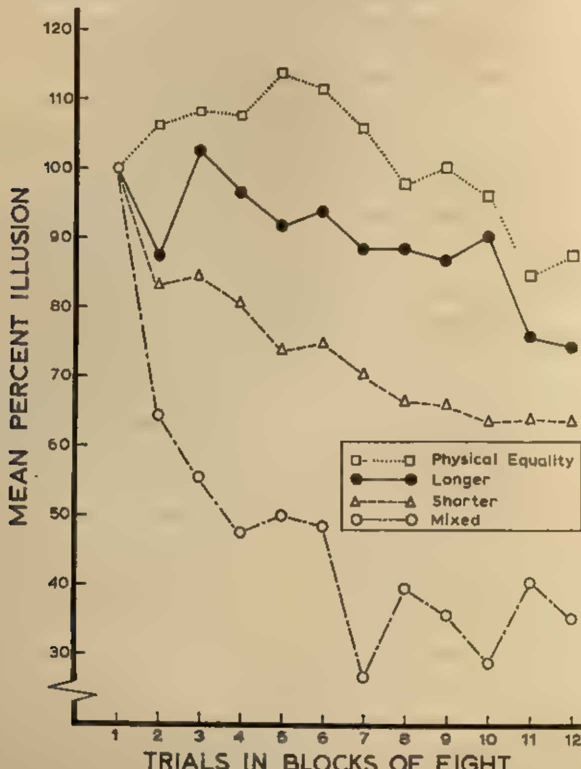

FIG. 1. MEAN PERCENTAGE THE ILLUSION OCCURRED IN THE FOUR PROCEDURES

The illusion in the second and successive blocks of trials is computed as a percentage of the amount of illusion on the first block of trials.

effect of trials was obtained for the mixed procedure and for the one where the variable was always set shorter than the standard, but not for the other two. Again, this finding is consistent with the visual picture given by Fig. 1. For the procedure where the variable was always set longer than the standard, and for the one where it was always set to physical equality, some decrement seems to occur for later, but not for earlier, blocks of trials.

Discussion

The evidence provided by this experiment appears to offer clear support for the conclusion that decrement in the *M-L* illusion with repeated observation is in part a function of the psychophysical procedure employed. The course of the illusion over trials is strikingly different for the different procedures.

The most probable interpretation of the superiority of 'mixed' procedure and of the 'shorter' procedure over the other two would seem to be in terms of the opportunities they provide for *S* to compare the two horizontal extents. With the mixed procedure, where prior to the *S*'s adjustment on successive trials the variable is set either obviously shorter or obviously longer than the standard, conditions for learning to occur would seem to be optimal. With the procedure where the variable is always set shorter than the standard, the illusion is absent prior to the *S*'s adjustment, since the variable is the line which apperas longer when it is physically equal to the standard, and this also seems to facilitate learning. In the case of the procedures where the variable is always set longer, or to physical equality with the standard, learning progresses more slowly if it occurs at all.

TABLE I

Analysis of Variance of the Data for the Four Procedural Groups

Source	*df*	*MS*	*F*	*p*
Between *S*s:	31			
procedures	3	56,664.16	2.58	N.S.
error (b)	28	21,912.33		
Within *S*s	320			
trials	10	2,040.13	32.61	<0.01
trials×procedures	30	253.48	4.05	<0.01
error (w)	280	62.56		

These two conditions appear to offer least opportunity for comparison of the horizontal extents under varied conditions.

Implicit in the preceding is the assumption that the decrement in the *M-L* illusion over trials is due to learning, rather than to a process of satiation, as Köhler and Fishback conclude.[11] The evidence for a learning interpretation appears now to be substantial and has been reviewed elsewhere so that a repetition would seem unnecessary.[12] With reference to the present experiment it is somewhat difficult to see how the various procedures employed would affect the decrement if satiation were the mechanism responsible. Presumably for satiation to occur, continuous or repeated presence of the *M-L* figure is all that is required, a condition which is met in all procedures. If, however, learning is what brings about the decrement, maximizing the opportunities for comparing the horizontal extents of the figure might be expected to facilitate the process. It would seem that randomly intermixing longer and shorter settings of the variable prior to *S*'s adjustment, and the use of the shorter settings alone, both exert this facilitative effect.

Summary

In this experiment decrement to the *M-L* illusion over trials was investigated as a function of psychophysical procedure. Four procedures

[11] Köhler and Fishback, *op. cit.*, 280.
[12] Parker and Newbigging, *op. cit.*, 134-135.

were compared. In the first, the typical procedure employed with the Method of Average Error, shorter and longer settings of the variable (the extent between the 'open' angles) were randomly intermixed. In the second only the shorter of the settings used in the mixed procedure were employed, while in the third the longer settings were used. In the fourth and final procedure the variable was always set to physical equality prior to *S*'s adjustment. The results showed that only the first two procedures noted produced a significant decrement to the illusion over trials. Since the last two procedures had been used in studies in which the decrement was not obtained under a variety of conditions, it is concluded that the failure is probably attributable to the psychophysical procedure employed. It is further concluded that the greater effectiveness of the mixed procedure in bringing about the decrement may be attributed to the more varied conditions it provides for the comparison of the horizontal extents, and learning is therefore facilitated.

MEMORY-SPAN WITH EFFICIENT CODING PROCEDURES

By Irwin Pollack and Lawrence B. Johnson
Air Force Systems Command, Bedford, Mass.

In 1952, Smith demonstrated that the memory-span for binary digits may be substantially increased following instructions about efficient coding procedures.[1] The present study is a direct extension of Smith's study to determine the conditions under which coding procedures are most effective. In particular, we investigated the reproduction of material presented verbally and in a tachistoscopic display. We also tested varying rates of presentation of verbal material.

Method and Procedure

Subjects. A group of 6 university students was tested 1.5 hr. per day on 28 successive testing days with both verbal and tachistoscopic materials. They were then taught the decimal equivalent of the 4-digit binary code for 12 days, and, finally, were retested for 10 additional days.[2]

Memory-span materials. Verbal messages, consisting of randomly selected scramblings of 'ones' and 'zeros,' were recorded. The 'zero' was read as 'oh.' Half of the messages were read without interruption; half of the messages, called "split messages," were read with a unit-pause after every four digits. The rate of reading was so adjusted that the average rate of presentation was constant at 1.4 binary digits per sec. S's task was to reproduce the message upon termination. The termination was indicated by a tonal signal. The answer sheet was subdivided into groups of four units. The length and 'split' of the message was announced before each presentation. Ten to 15 different messages of each message-length were recorded to reduce the learning of specific messages.

Tachistoscopic materials. The tachistoscopic messages were 8 circles arranged in a horizontal line, exposed for 0.04 sec., subtending a visual angle of 11° at a seating distance of 16 ft. Each circle was either filled or unfilled. A pool of over 100 dis-

* Received for publication November 23, 1962. Technical Documentary Report ESD-TDR-62 of Electronic Systems Division, Project 7682.

[1] S. L. Smith and G. A. Miller. The effects of coding procedures on learning and memory. *Res. Lab. Electronics, Mass. Institute of Technology. Quart. Progress Rept.*, Dec. 16, 1952. This work is also cited by G. A. Miller, The magical number seven, plus or minus two: Some limits on our capacity for processing information, *Psychol. Rev.* 63, 1956, 81-97.

[2] The initial experimental design called for two groups: an experimental group and a control group. The control group's treatment was identical, except that the intermediate coding instruction was omitted. Under the group testing, which we employed, the control group noted the superiority of the experimental group and defeated their purpose by learning the code clandestinely. Only the results of the experimental group ($N = 6$) are considered, except when noted.

plays was scrambled on successive days. In half of the slides, a dividing line split the display into two groups of four units each. Twenty-four 'split' and 24 unbroken displays were presented on each testing day.

Coding instruction. Instruction about coding was accomplished by explaining the principle of binary numerals. The following code was employed: 0000 to 0; 0001 to 1; . . . , and 1111 to 15. *Ss* were scored with flash cards under conditions of timed trials.

After instruction about coding, *Ss* were instructed to employ the shortened notation. We shall refer to the notation as decimal notation, although the paired numbers 10, 11, . . . 15 were scored as a single number. The notation was soon abandoned by *Ss* with tachistoscopic test-materials for the *Ss* reported that the decimal

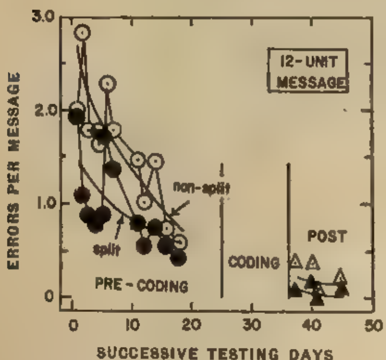

Fig. 1. Improvement in Memory-Scores with Practice and with Instruction about Coding
(The ordinate is the average number of binary digits incorrectly reproduced. Each point is based upon 5 messages to each of 6 *Ss*.)

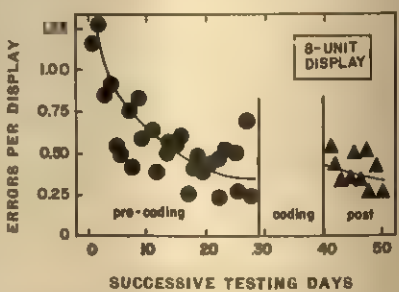

Fig. 2. Improvement in Tachistoscopic Recognition with Practice
(The ordinate is the average number of binary digits incorrectly reproduced. Each point is based upon 23 presentations to each of 6 *Ss*.)

responses were made only after they had already recorded the binary responses. On the other hand, decimal notation was enthusiastically employed with the verbal materials.

Scoring. An error in decimal notation was treated as equivalent to a response of four binary digits. Such an error will be termed a "sectional error." Reconversion to binary errors yields 2.13 binary errors per error in decimal notation on the assumption that all errors in decimal were equally-likely. This method was conservative; sample comparisons between errors in decimal with the original binary messages yielded 1.9–2.0 binary errors per error in decimal notation.

Results

Learning the binary code. The average time taken by *Ss* to make the 4-digit binary to decimal conversion decreased from 6 sec. per item on the 1st trial, to 3 sec. on the 5th trial, and to 1 sec. per item by the 30th trial.

Effect of practice and coding instruction. Fig. 1 presents the improvement in reproducing 12-unit binary messages as a function of the number of successive testing days. Substantial improvements in performance are obtained without specific coding instruction. Following such instruction, there is a further improvement in performance.

Fig. 2 presents corresponding results with the tachistoscopic materials for the split display. (Equivalent results were obtained for unbroken displays.) Again, substantial improvements in performance are obtained early

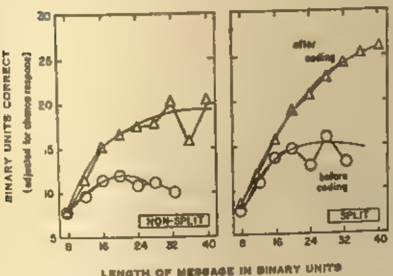

FIG. 3. MEMORY SCORES BEFORE AND AFTER CODING INSTRUCTION
(The right section represents messages in which a pause is interposed between successive groups of four binary digits. The left section represents unbroken messages. Each point is based upon 18 messages presented to each of 6 *Ss*.)

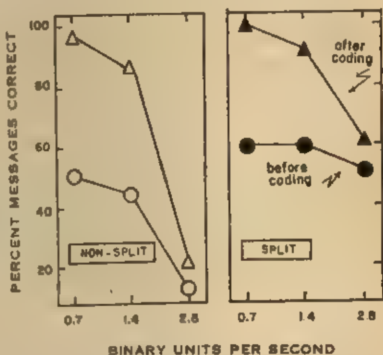

FIG. 4. EFFECT OF CODING INSTRUCTION FOR VARIOUS RATES OF MESSAGE PRESENTATION
(See footnote of Fig. 3. Each point represents the results of 5 messages presented to each of 6 *Ss*.)

in practice without specific coding instruction. Unlike the verbal messages, however, there is little systematic gain following coding instruction.

Effect of coding on the reproduction of verbal messages. The number of binary units correctly received before and after coding instruction, is presented in Fig 3 as a function of the length of the messages in binary units. The left section of the figure describes the results of the non-split or ungrouped messages. The right section represents the results for messages split into sections of four binary units each. Because of the large initial changes in performance, only the messages presented during the last seven days prior to coding are considered. The number of binary units achieved after coding instruction is nearly double that achieved after approximately three weeks of practice. A comparison of the results from the initial stages of pre-coding with the post-coding results would be even more impressive. In all cases, there were substantial gains following the coding instruction

for each S. The better Ss were often able to receive the split 40-unit messages without error.

Effect of rate of presentation. Recoding of binary digits into decimal digits takes time and therefore might be expected to be relatively ineffective at rapid rates of presentation. To test this point, the Ss were presented with 16-unit binary messages at three rates of presentation, before and after coding instruction.

The results of these tests are presented in Fig. 4. Large differences are

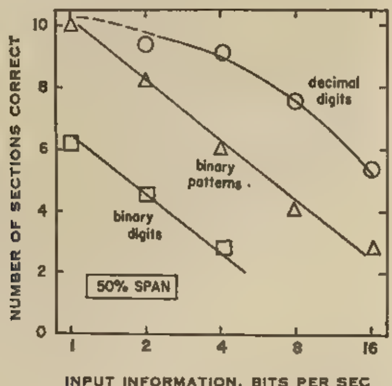

FIG. 5. EFFECT OF FORMAT OF VISUAL PRESENTATION ON MEMORY-SPAN
(The abscissa is scaled in terms of the number of bits per second. In terms of the number of flashes per second, the abscissa should be divided by 4.0 for the two upper curves. Each point is based upon 9 messages presented to each of 12 Ss.)

observed between the pre- and post-coding instruction with non-split and split messages at slow rates of presentation; however, these differences are sharply attenuated at the faster rates of presentation.

Effect of visual format. Additional tests were carried out with the combined experimental and control groups to examine the effect of visual format. Three formats were employed. Decimal digits (0 to 15) were programmed by a 4×4 display, lighting the digits 0 to 15. Binary digits were programmed upon the same apparatus by utilizing only the digits 0 and 1. Binary patterns were represented by a line of four small bulbs, arranged horizontally. The decimal response was employed for all tests.

The results of these tests are considered in Fig. 5 in terms of the num-

ber of sections, or decimal digits, at which 50% correct reproduction is achieved. The abscissa represents the rate of presentation of the numerals in terms of the input transmission in bits per second. At slow rates of presentation, there is little difference between the presentation of decimal digits and binary patterns. At faster rates, however, the decimal digits are clearly superior. The 4-digit binary patterns are consistently superior to the binary digit-by-digit sequence which flashed at four times the rate of the pattern display for equivalent input rates.

Effect of auditory format. A series of messages containing 32 binary units (split and non-split) was programmed to note the effects upon recall when the method of reading the series was varied. The binary digits were read in groups of 1, 2 and 4 digits, *e.g.* 1001 was read as one, zero, zero, one; or ten, one; or a thousand and one. Decimal digits from 0 to 15 were also presented.

Significant gains ($0.02 < p < 0.01$) were obtained after coding for formats of 1 and 4 binary digits; as well as split messages with a format of 2 binary digits (0.05 level). No differences were observed after coding with the decimal digits. All tests employed the Wilcoxon Groups of Replicates Test.[3]

Discussion

The present study confirms Smith's findings that instruction with efficient coding procedures will result in larger memory-spans. We further show that the effect of coding instruction is modified by the rate and format of presentation. We demonstrate that coding procedures are effective primarily at slow rates of message presentation, and that coding procedures are most effective when the format of message presentations is compatible with the coding. Both of these findings are reasonable in terms of the temporal and organizational requirements for coding.[4]

Despite the large gains in memory-span, we failed to obtain a 4-fold improvement in memory-span with the use of a 4-binary digit-code. Our failure is related to the fact that the later pre-coding scores had demonstrated substantial improvements over the initial test-scores. Presumably our *S*s were already learning to impose an inefficient coding scheme. Their verbal reports confirm this approach. Also, the decimal code employed (0 to 15) invites redesign. All of the *S*s volunteered statements that performance would have been better had they learned 16 individual symbols rather than combinations of decimal digits.

[3] F. Wilcoxon, Some rapid approximate statistical procedures, American Cyannamid Co., Stamford, Connecticut, 1949.

[4] Miller, Human memory and the storage of information, *IRE Transactions on Information Theory,* Vol. IT-2, No. 3, Sept. 1956, 129-137.

The failure of our *S*s to use the coding scheme with tachistoscopic materials is consistent with the relative ineffectiveness of the coding procedures at rapid rates of presentation. The temporal requirements of the task do not permit the fullest utilization of efficient coding procedures at rapid rates of presentation.[5]

Summary

Binary verbal messages and tachistoscopic materials were presented before, and after, instruction with a binary to decimal digit-code. Confirming Smith, we find substantial improvements in memory-span for verbal messages after coding. We also find little change for tachistoscopic materials. The effect of coding is enhanced by compatible message formats and at slow rates of presentation. Effective utilization of efficient coding procedures requires sufficient time for recording operations.

[5] Klemmer demonstrated little or no improvement in the reporting of tachistoscopically flashed binary patterns when recorded in octal digits. (E. T. Klemmer. Does recording from binary to octal improve the perception of binary patterns? *J. exp. Psychol.*, 67, 1964, 19-21.)

A REPLICATION OF THE MEL SCALE OF PITCH

By Robert J. Siegel, Queens College

Stevens, Volkmann, and Newman derived a scale measuring the subjective pitch of tones by the method of bisection.[1] They called the subjective unit of pitch, the "mel" and arbitrarily assigned a value of 1000 mels to the pitch of a 1000 ~ tone. According to their investigation, a given frequency-interval (on the logarithmic scale of music) would appear smaller in size at the upper and lower ends of the frequency-scale than in the central region; this tendency being more marked in the lower than in the upper region.

More recently, two replications of Stevens, Volkmann, and Newman's scale have been reported. The first was performed by Stevens and Volkmann by the methods of equal sense distances and again by bisection. These results were essentially the same as those of the original experiment.[2]

The second replication was performed by Lewis using the methods of bisection, fractionation, and equation of supraliminal extents of pitch.[3] He did not find satisfactory agreement between his several scales and those derived by Stevens and his associates, and attributed the discrepancies to variations in *O*s and in the particular measure used. Unfortunately, only an abstract of this paper was published and the exact details of Lewis' results are not known.

In view of these conflicting reports and the lack of recent work on this subject, the present study was undertaken.

Observers. Ten undergraduate students, three women and seven men, all under 25 yr. of age, with mean of about 20 yr., took part in the experiment. None of them had any musical training or any obvious hearing defects.

Apparatus. The basic apparatus was an Eico model 377 audio generator and a 6-in. speaker in a bass-reflex type enclosure. The audio generator was calibrated by comparison with the standard 60 ~ alternating line current and its harmonics.

Procedure. O was seated in a chair with the speaker-enclosure raised to the level of his head and placed about 1 ft. behind him. *E* was seated across a small table from *O* at approximately the same distance from the speaker. The audio generator was placed on the table in such a manner that *O* could not see its con-

* Received for publication Sept. 24, 1963. The author wishes to acknowledge the assistance of Dr. William F. Reynolds of Queens College in the preparation of this article.

[1] S. S. Stevens, John Volkmann, and E. B. Newman, A scale for the measurement of the psychological magnitude of pitch, *J. acous. Soc. Amer.* 8, 1937, 185-190.

[2] S. S. Stevens, and John Volkmann, The relation of pitch to frequency, a revised scale, this Journal, 53, 1940, 329-353.

[3] Don Lewis, Pitch-scales. *J. acous. Soc. Amer.*, 14, 1942, 127, (abs.).

trol knobs and dials. *O* was instructed that a standard tone would be sounded and would be alternated with a lower tone at 2-sec. intervals, and that his task was to instruct *E* how to adjust the pitch of the lower tone until it sounded exactly half as high as the upper (standard) tone. The point of bisection was read on the dial of the audio generator and noted. Many of the *O*s expressed concern that they would make "incorrect" judgments, but were told not to worry and simply to do their best. Each bisection took approximately 1 min. One practice-trial was given with a 1500~ tone as the standard.

Tones of each of the following frequencies were bisected twice by each *O*: 92~; 165~; 510~; 920~; 1650~; 5100~; 9200~; this made a total of 14 bisections. Loudness of tones was maintained at approximately 50 db. (as heard by *O*) by minor adjustments by *E* of the loudness control of the audio generator. The order of presentation of the 14 tones was random with a different order for each *O*. The initial "lower tone" for each standard tone was randomly selected within a range of three-quarters of an octave above and below the tone of one-half the

TABLE I

GEOMETRIC MEANS AND MEASURES OF VARIABILITY OF FREQUENCIES JUDGED HALF THE PITCH OF SEVEN STANDARD FREQUENCIES

(Every one of the 10 *O*s made two judgments)

Standard Frequency:	92	165	510	920	1650	5100	9200
Frequency Judged Half:	54.9	88.9	254	424	793	2375	3990
Inter-*O* Variability	9.6%	6.5%	2.8%	6.4%	6.7%	3.3%	17.4%
Intra-*O* Variability	6.6%	6.7%	2.4%	8.3%	8.7%	3.7%	7.6%

Average inter-*O* variability $(V)=7.5\%$.
Average intra-*O* variability $(V_o^{av})=6.3\%$.

frequency of the standard. In no case was *O* satisfied that any initial "lower tone" was exactly one-half the pitch of the standard tone.

Results and discussion. The geometric means of all 20 bisections of each standard were calculated and are presented in Table I together with the inter-*O* variability (V) and the average intra-*O* variability (V_o^{av}) for each standard tone. (These measures of variability are the same as were used by Stevens and his associates and are used here in order that the variabilities of the two experiments can be more easily compared.) The frequencies of the standards are plotted against the geometric means of the frequencies of the bisection (frequencies judged half) together with Stevens, Volkmann, and Newman's corresponding data in Fig. 1.[4] The mel-curves derived from both the present data and from Stevens, Volkmann and Newman's data are presented in Fig. 2. The method of derivation used was the same for both sets of data.[5] It may be seen from Fig. 1 that

[4] Stevens, Volkmann and Newman, *op. cit.*, 187.

[5] A value of a 1000 mels is assigned to the pitch of a 1000 ~ tone. The frequency of the tone which would have been judged half the pitch of the 1000 ~ tone (found by projection on the data-curve, Fig. 1) is assigned a pitch of 500

the curve taken from the data of the present study is a straight line while Stevens, Volkmann, and Newman's corresponding curve is a somewhat meandering slightly curved line. The mel-curve derived in the present study is an exponential curve (similar to the curve for subjective *loudness* of tones) while Stevens, Volkmann and Newman's curve tends to be ogival.[6]

Considering the subjective nature of the judgments involved, the re-

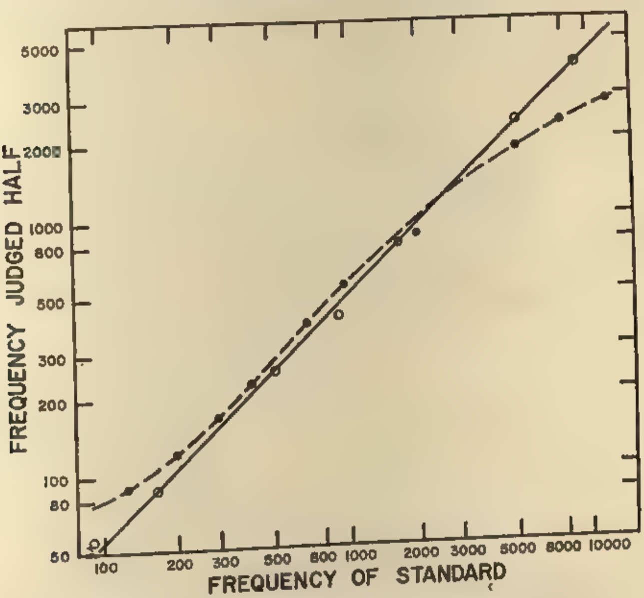

Fig. 1. Frequency of Standard Tone vs. Frequency Judged Half as High

The solid curve represents the data obtained in the present experiment; the broken curve represents the data of the 1937 experiment of Stevens, Volkmann, and Newman.

sults are reasonably consistent in both this study and in the studies of Stevens and his associates (the exact variability measures will be presented later) so that we may argue that the source of the discrepancies between the two sets of data would not likely be due to chance; the dis-

mels. The frequency which would have been judged half the pitch of the 500-mel frequency is assigned a pitch of 250 mels. The process is repeated for 125 mels. For values above 1000 mels the process is reversed. Stevens applies this method of derivation graphically, but because the data here were represented by a straight line, the method was applied mathematically.

[6] S. S. Stevens and H. Davis, Psychophysiological acoustics, pitch and loudness, *J. acous. Soc. Amer.*, 8, 1936, 1-13.

crepancies are more likely due to some actual error in apparatus or procedure which caused either the pitch of the higher tones used in Stevens' experiments to be perceived as lower than they actually were, or the pitch

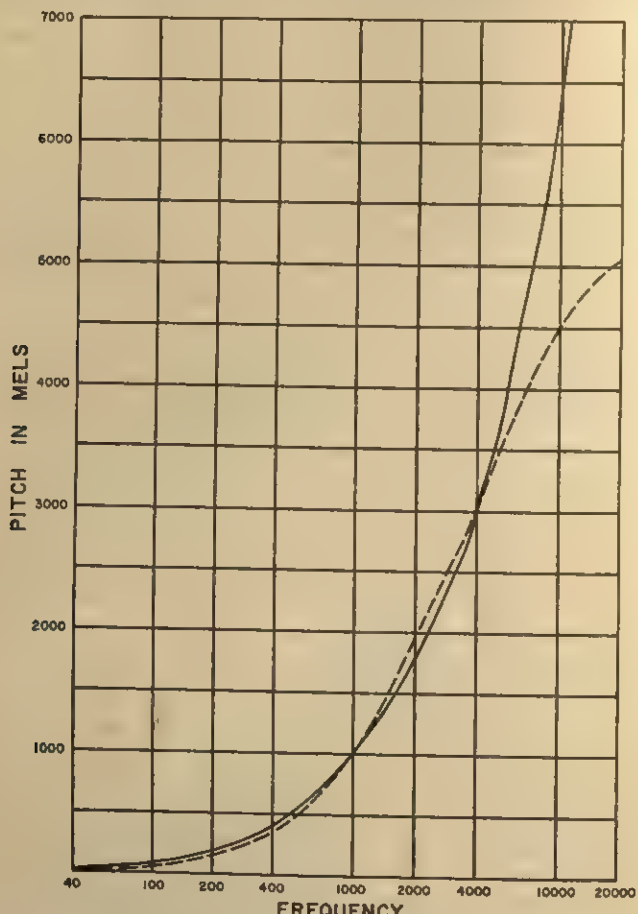

FIG. 2. FREQUENCY VS. PITCH IN MELS
The solid curve represents the data obtained in the present experiment; the broken curve represents the data of the 1937 experiment of Stevens, Volkmann, and Newman.

of the higher tones used here to be perceived as higher than they actually were. Examination of Stevens' experiments reveals two factors which could have caused the required effect: the room(s) in which the experiments were conducted, and the ages of the *O*s.

Pitch, being a subjective dimension, can depend on many variables other than frequency, and an experiment attempting to equate pitch with frequency must seek to eliminate as many of these extraneous variables as possible. In Stevens' experiments, the walls of the room(s) used were made highly acoustically absorbent by being lined with rockwool; this practice has a high degree of face validity, but when the natures of sound-propagation and absorption are examined, the procedure is revealed to be highly questionable. They write:

> With absorbing materials for which the absorption depends primarily on the viscous resistance encountered in small channels or interstices, we must expect that the absorption will increase with the frequency of the sound, so that high notes will suffer a relatively greater absorption than low notes.[7]

Rockwool is such a material.[8]

If the source of sound is at a relatively great distance from *O*, as it was in Stevens' experiments (1937: at the end of the room; 1940: one at 7 ft. and one at 3 ft.), the nature of the room will be a potent factor affecting the perceived pitch of tones.[9] What this means specifically in terms of Stevens' experiments is, because the energy of the low tones was reflected from the room-surfaces back to *O* to a greater degree than from the high tones: (1) that the proportion of energy reaching *O*'s ears from the low tones was greater than from the high tones, and (2) that the low tones would be perceived to have a greater spatial extent due to the presence of their echoes coming from all directions, while the high tones would be perceived to emanate from almost a point-source due to their conspicuous lack of echoes.

Both these factors would have served to lower the perceived pitch of the higher tones in Stevens' experiments.[10] In the experiment performed here, *O* was only one foot from the speaker, thus minimizing the relative intensities of the room-echoes. Furthermore, the experiment was conducted in a small room with painted metal walls which tended to make any echoes produced have fairly constant effects at all frequencies of sound.

The ages of *O*s could also have caused an apparent effect of the lowering of the pitch of the higher tones in Stevens' experiments. Although no details about the *O*'s were noted, these may be inferred: (1) in the 1937 experiment, Volkmann and Newman were *O*s being about 30 and 28 yr. old, respectively; (2) in the 1940 experiment, Stevens and Volkmann were *O*s, both being about 34 yr. old. It was reported that the other *O*s in both experiments were graduate students, pre-

---

[7] G. W. Stewart and R. B. Lindsay, *Acoustics*, 1930, 300-301.

[8] H. F. Olson and F. Massa (*Applied Acoustics*, 2nd ed., 1939, 385) gives the absorption coefficients for 1 in. thick rockwool as: 0.35 at 128 ~; 0.49 at 256 ~; 0.63 at 512 ~; 0.80 at 1024 ~; 0.83 at 2048 ~.

[9] For a basic discussion of the effects of room acoustics, see E. T. Canby, *Home Music Systems*, rev. ed., 1955, 193-209.

[10] At this point we may cite a common example of such phenomenon in order to make these concepts more concrete In an acoustically highly reflective room such as a tile bathroom, sounds appear unusually crisp and clear to the noticeable presence of echoes, especially of the higher components of the sounds produced therein; in an acoustically highly absorbent room, such as a closet full of clothing, or a heavily draped room with many rugs, sounds appear unusually dull and muffled, due to the noticeable lack of echoes, especially of the higher components of the sounds.

sumably most of whom were over 20, and possibly some significantly older than 20. Over the age of 20 yr., hearing begins to lose a great deal of sensitivity, especially in the higher ranges.[11] If, in fact, the ages of some *O*s were significantly over 20 yr., we may expect to find an apparent lowering of the pitch of the higher tones occuring jointly with an increase of the variabilities of judgements. Both phenomena were observed.[12] In the present experiment, the ages of all the *O*s were lower than 25, with mean age estimated to be lower than 20 yr.

Because of these facts, the mel-curve derived in this experiment appears to be more valid than the curves obtained by Stevens, Volkmann, and Newman, and by Stevens and Volkmann. The present finding indicates that the mel-curve of subjective pitch vs. frequency to be of the same form as the *sone* curve of subjective loudness, that is, an exponential curve.

---

[11] J. C. R. Licklider, Basic correlates of the auditory stimulus, in S. S. Stevens (ed.) *Handbook of Psychology,* 1951, 996-997.

[12] 1937; $V = 12.7\%$, $V_o^{av} = 11.7\%$; 1940 fractionation: $V = 9.4\%$, $V_o^{av} = 15.6\%$; the present experiment: $V = 7.5\%$, $V_o^{av} = 6.3\%$. The V_o^{av} for the 1937 experiment was only given for two of the 5 *O*s.

ISOLATION-EFFECTS DURING PAIRED-ASSOCIATE TRAINING

By Slater E. Newman
The University of North Carolina at Raleigh

These experiments were done to determine whether isolating the stimulus- or the response-term of a paired associate has the same effect on learning as isolating an item in a serial-list. Two techniques for producing isolation were used, color and low meaningfulness. In serial learning, each has been found to facilitate learning of associations involving the isolated term as a response and as the nominal stimulus.[1]

In paired-associate learning, isolating the stimulus-term and the response-term has been found to facilitate learning the isolated pair.[2] The techniques used to produce isolation in these experiments differed from those used in serial learning studies. Erickson, however, has shown that when one or both terms of already isolated pairs were further isolated (through color), learning of the pair was facilitated, while Newman and Forsyth have reported that when paired associates were presented serially, isolating the stimulus-term or the response-term through color was facilitative.[3]

Method: Subjects. Each group had 15 *Ss*. For Experiments I ($N = 45$), II ($N = 75$), and III ($N = 75$), the *Ss* were undergraduates. For Experiments IV ($N = 60$) and V ($N = 60$), The *Ss* were nurses-in-training at Dorothea Dix State Hospital. In experiments in which isolation was produced through color, every *S* was tested for color-blindness at the end of the session. Only data from the first 15 noncolor-blind *Ss* of a group were analyzed.

---

* Received for publication December 21, 1964. Support for this research was provided by grants from the North Carolina State Faculty Research and Professional Development Fund and by the Office of Naval Research under contract Nonr 486(08). This is Technical Report Number 12 prepared under this contract.

[1] F. N. Jones and M. H. Jones, Vividness as a factor in learning lists of nonsense syllables, this JOURNAL, 55, 1942, 96-101; S. E. Newman and Eli Saltz, Isolation effects: Stimulus and response generalization as explanatory concepts, *J. exp. Psychol.*, 55, 1958, 467-472; H. E. Rosen, D. H. Richardson, and Eli Saltz, Supplementary report: Meaningfulness as a differentiation variable in the von Restorff effect, *ibid.*, 64, 1962, 327-328; M. H. Smith, Jr., and Ellen Stearns, The influence of isolation on the learning of surrounding materials, this JOURNAL, 62, 1949, 369-381.

[2] R. L. Erickson, Relational isolation as a means of producing the von Restorff effect in paired-associate learning, *J. exp. Psychol.*, 66, 1963, 111-119; Jacob Nachmias, Henry Gleitman, and V. V. McKenna, The effect of isolation of stimuli and responses in paired associates, this JOURNAL, 74, 1961, 452-456.

[3] Erickson, *op. cit.*, 114-116; S. E. Newman and G. A. Forsyth, Supplementary report: Isolation effects when paired associates are presented serially, *J. exp. Psychol., in press.*

Materials. The list was composed of 14 paired associates, dissyllables from Noble's list, each with its accent on the first syllable.[4] The mean 'm'-value for the stimulus-terms was 6.70 and for the response-terms, 6.67; the medians were 7.15 and 7.14. The range for the stimulus-terms was 3.21 to 9.43, and for the response-terms, 3.36 to 9.61. No two stimulus-terms or response-terms began with the same letter. No obvious associations existed between the terms of a pair. When isolation was produced through color (Experiments I, II, and III), the isolated term of the pair was presented in red. When isolation was produced through low meaningfulness (Experiments II, III, IV, and V), a term with an 'm'-value less than 1.29 was inserted in place of one of the terms of the critical pair. The critical pairs in these experiments were chosen from those which, in previous experiments with this list, were of greater than median difficulty.[5]

Procedure. The items were presented on a Lafayette memory-drum. A pairing-test procedure was used, every pairing-trial being followed by a test. A 2-sec. rate was used for both pairing- and test-trials. The intertrial-interval was 4 sec. Stimulus-terms were isolated on both pairing- and test-trials; response-terms were isolated only on pairing-trials. Three different pairing-orders and three different test-orders were used to reduce the likelihood that serial-position would serve as a cue. Each *S* was given 15 trials to learn the list. In each experiment, a control group learned a list on which none of the items was isolated.

In Experiment III, training was followed by two R-S tests. A different order was used on each test. The rate was 2 sec. All items were typed in black. In Experiments IV and V, training was followed by a recall of the response-terms and then by a recall of the stimulus-terms. *Ss* were given 45 sec. for each.

Results. The number of correct responses for the critical pair during 15 trials was used to measure the *absolute* effect of isolation. In every experiment, *t*-tests were made comparing the mean of each group with the mean for the control group. These means and the results of the *t*-tests are presented in Table I. Since there were several independent tests for each of the comparisons, the tests were combined for each comparison.[6] These results showed that, when the stimulus-term was isolated through color or low 'm,' facilitation occurred ($P < 0.01$). When the response-term was isolated through color, facilitation was slight ($0.30 > P > 0.20$), but when isolation was produced through low 'm,' learning was retarded ($P < 0.01$).[7]

The rate of learning the critical pair contrasted with other pairs in the list is designated as the *relative* effect of isolation. This was determined

[4] C. E. Noble, An analysis of meaning. *Psychol. Rev.*, 59, 1952, 421-430.

[5] In serial learning, the critical items have also been of greater than median difficulty, perhaps due to their being from the middle of the list where learning is slow.

[6] B. J. Winer, *Statistical Principles in Experimental Design*, 1962, 43-44.

[7] For number correct, two of the four comparisons where learning was retarded were significant ($p < 0.05$); for rank, three were significant ($p < 0.05$). The one comparison for each variable in which facilitation occurred was not significant —($0.70 > p > 0.60$) for number correct, ($0.30 > p > 0.20$) for rank.

for each S by ranking the number of times he got each term correct during training. Comparison of the mean rank for the critical pairs gave results similar to those for number correct. Facilitation occurred when the

TABLE I

PERFORMANCE DURING AND AFTER TRAINING FOR EACH EXPERIMENT

Experiment	Condition	Mean correct (Critical pair)[a]	Mean rank for number correct (Critical pair)[a,b]	Median trial on which R-term for critical pair was first given[c]	Mean correct during training (All pairs)[a]	Number correct on post-training Tests (Critical pair)[d,e]
I	Control	7.80	9.90	6.0	139.07	
	S-Color	11.53†	6.57†	3.0	148.73	
	R-Color	9.13	7.00*	5.0	120.20	
II	Control	8.20	8.93	4.0	128.00	
	S-Color	12.07†	3.57†	3.0	127.20	
	R-Color	9.67	7.20	3.0	138.13	
	S-'m'	12.73†	4.63†	2.0	149.87	
	R-'m'	5.40	11.77†	8.0*	129.13	
III	Control	9.00	8.83	4.0	136.53	9, 8
	S-Color	11.20	5.83*	3.0	133.40	9, 10
	R-Color	11.20	5.33*	3.0	133.33	10, 8
	S-'m'	9.67	8.00	4.0	131.73	5, 5
	R-'m'	9.87	6.93	3.0	132.00	12, 12
IVa	Control	7.73	9.93	2.0	141.33	8, 11
	R-'m'	4.00*	13.57†	9.0†	139.47	6, 6
IVb	Control	11.80	4.90	2.0	141.33	10, 11
	R-'m'	8.13*	10.83†	3.0	148.33	11, 12
IVc	Control	10.13	7.90	4.0	141.33	7, 9
	R-'m'	7.73	9.77	6.0	136.80	8, 8
Va	Control	8.07	9.70	3.0	142.53	11, 6
	S-'m'	11.00	7.07	3.0	138.73	10, 6
Vb	Control	10.53	7.07	3.0	142.53	10, 9
	S-'m'	12.60	4.70*	4.0	140.27	8, 7
Vc	Control	9.13	8.90	4.0	142.53	10, 7
	S-'m'	9.80	7.70	3.0	137.73	7, 7

* $p<0.05$. † $p<0.01$.

[a] The mean of each group was compared with that for the control group, using a t-test.

[b] The pair learned fastest was ranked 1.0.

[c] A Kruskal-Wallis test was used for each comparison.

[d] Chi-square was used for each comparison.

[e] In Experiment III, two R-S tests were given; in Experiments IV and V, R-term recall, then S-term recall. No post-training tests were given in Experiments I and II.

stimulus-term was isolated through color or low 'm' and when the response-term was isolated through color ($P < 0.01$). Learning was retarded when the response-term was isolated through low 'm' ($P < 0.01$).

To determine whether isolation affected the rate at which the list was

learned, the trial on which the entire list was given correctly was recorded for each *S* and a Kruskal-Wallis test done on these data for comparisons in every experiment. None of the *H*'s was significant ($P > 0.05$). Analyses of variance for total correct during training also gave no significant *F*s in any experiment ($P > 0.05$).

The effect of isolation on learning of response-terms was studied by determining the trial on which the response-terms were first given. Combined probabilities from Kruskal-Wallis tests revealed that when isolation was produced through low 'm,' the response-term of the critical pair tended to occur later during training than when it was isolated ($P < 0.05$).[8] No other comparisons were significant ($P > 0.05$).

Associative learning was studied by determining the percentage of *S*s who gave the response-term to its correct stimulus-term the first time they gave it. Isolation had no effect. Most *S*s gave the response-term correctly the first time.

Two types of error were studied: (1) the number of times that the response-term from the critical pair was given in place of other response-terms from the list; and (2) the number of times that other response-terms from the list were given in place of the response-term for the critical pair. For each *S*, both scores were divided by the total number of errors involving response-terms from the list and Kruskal-Wallis tests performed on each of these variables. None of the *H*'s was significant ($P > 0.05$).

The total number of times the stimulus-term from the critical pair was given correctly during each R-S test was determined for Experiment III. Isolation had no effect ($P > 0.05$). Neither did it affect performance on recall of the R- or S-terms during Experiments IV and V. Neither total recall, not number of items recalled from the critical pair, differed due to isolation of stimulus- or response-terms by color or 'm' ($p > 0.05$).

Discussion. Both variables, number correct and rank, permit the same conclusions. When the stimulus-term was isolated, performance was facilitated; when the response-term was isolated, there was a tendency for facilitation to occur when isolation was produced through color. When, however, isolation was produced through low meaningfulness, learning was retarded.

Only when the response-term was isolated through low meaningfulness

[8] Two of the four comparisons where learning was retarded were significant ($p < 0.05$). The one comparison in which facilitation occurred was not ($0.50 > p > 0.40$).

do these results differ from those for serial learning.[9] In the present experiments learning was retarded, while in serial learning, facilitation occurred.

When isolation of the response-term is produced through low meaningfulness, two effects may be involved: facilitation due to isolation, and retardation due to low meaningfulness. The net effect of isolation produced through this technique will depend on which of these effects is stronger. The data for response-term learning combined with the slower learning of the pairs isolated through low 'm' suggest that in these experiments the retardation resulting from low meaningfulness exceeded the anticipated facilitation due to isolation. Horowitz has proposed a similar analysis of the effects of isolation of the response-term produced through low 'm.'[10]

Why then is there facilitation during serial learning? The disparity may be a function of differences between serial and paired-associate tasks. When items appear in the same order on each trial, or when most of the items serve as both a nominal stimulus and as a response-term, the facilitation due to isolation may exceed the retardation resulting from low meaningfulness. Thus, facilitation would be observed.

The disparity might also be a function of procedural differences between the serial and paired-associate experiments. In the Newman and Saltz' experiment, the *S*s pronounced each term in the list aloud during the first trial.[11] In the experiments on paired-associates, overt pronunciation of the response-terms was not required during any of the pairing trials. Thus, isolating the response-term through low meaningfulness may be facilitative only after the isolated term has been given overtly.

The number of times the stimulus-term of the critical pair was given in R-S recall (during Experiment III) and in S-term recall (during Experiments IV and V) did not differ as a function of whether the stimulus-term was isolated during training. This suggests that during training the type of response (*e.g.* pronunciation) to a stimulus-term isolated through color and through low meaningfulness did not differ from the type of response to stimulus-terms that were not isolated.

In these experiments, the critical pairs, when not isolated, were more difficult to learn than other pairs from the list. Under these conditions, facilitation was probably more likely than if the critical pairs had been

[9] Newman and Saltz, *op. cit.*, 468.
[10] L. M. Horowitz, Associative matching and intralist similarity, *Psychol. Rep.*, 10, 1962, 751-757.
[11] Newman and Saltz, *op. cit.*, 467-472.

chosen from the easier items. Erickson, however, has shown that isolation (at least when produced through color) is facilitative even for the easier pairs of a list.[12] Whether or not the effects of isolation are a function of the difficulty of the critical pair has not been systematically investigated. A similar question remains concerning the effects of isolation in serial learning.

Summary

When the stimulus-term of a paired associate was isolated through color or low meaningfulness, or its response term was isolated through color, learning of the pair was facilitated. When the response-term was isolated through low meaningfulness, learning of the pair was retarded.

---

[12] Erickson, *op. cit.*, 114-116.

INDIVIDUAL VARIATIONS IN THE POSTOCULAR LINES OF REGARD

By MARSHALL B. JONES, University of Florida

Ordinarily the visual, somesthetic, and gravitational frames of reference are aligned. It sometimes happens or can be arranged, however, that they are not, as when the head or body is tilted or the axes of the visual or gravitational fields are changed. Situations in which two perceptual frames of reference are at odds with each other have been frequently explored.[1] The general result is a compromise between them. Typically, however, there are pronounced individual variations. Some individuals, in a human centrifuge for example, locate the 'vertical' with the gravitational field; others locate it with the visual field; the majority locate it between the two. These variations have occasioned considerable interest, particularly in relation to personality.

The present study has to do with a situation in which visual and somesthetic frames of reference are discrepant. It differs from most of the situations which have been studied in that it occurs frequently in everyday life and requires no special apparatus of any consequence.

Subjectively, the human visual field is monocular and the one Cyclopean eye is located in the midline of the head. Objectively, of course, the eyes are centered 30 mm. or more laterally from the mesial line. In consequence (objectively) monocular vision involves a discrepancy between the visual field, which is offset some 30 mm. from the midline, and the somesthetic frame of reference with its major axis in the median plane.

To demonstrate this discrepancy it is only necessary to have *S* close one eye and look down a straight edge directly in front of his open eye. Then ask *S* to tell you when the touch of a pencil on his scalp seems to him to be lined up with the straight edge. Some people consistently locate the "postocular line of regard" in the midline of the head. Others locate it directly behind the open eye. A few feel that the touch is lined up when it is on the contralateral side of the head or past the optical axis on the ipsilateral side. Most people locate the postocular line between the midline and the optical axis. These judgments are easily made and quite definite.

* Received for publication October 29, 1964. The data reported here were collected at the U.S. Naval School of Aviation Medicine, Pensacola, Florida.

[1] H. A. Witkin, H. B. Lewis, Max Hertzman, Karen Machover, P. B. Meissner, and Seymour Wapner, *Personality through Perception*, 1954, pp. 24-76.

The purpose of this study is to determine the distribution of the post-ocular lines, their relation, if any, to hand- and eye-dominance, and to gain some notion of their determinants.

Procedure. Each *S* was seated at a table with his chin and forehead in a head-rest. A patch was placed over the eye not in use. A stand holding the straight edge was set up on the table 2 ft. from *S*. The straight edge (a ruler) was so adjusted that it stood directly in front of and at the same level as *S*'s open eye. When S looked down it, he saw only the forward end of the ruler.

For half the *Ss* the right and for the other half the left postocular line was determined first. *E* stood behind *S*, who was instructed to tell *E* at each touch of the pencil on his scalp whether the touch felt to the right, to the left, or lined up with the ruler. If *S* had difficulty in understanding, he was asked to imagine a line which passed along the ruler and came straight back to his head. He was to report whether the touch on his scalp lay on or to the right or left of this imaginary line. When *S* indicated that the touch was "on the line," the point was marked on the scalp. Then the succession of points was continued until it reached the midline.

E then returned by short steps from the midline to the right temple. When *S* indicated that the touch was on the line, this second point was marked, and *E* then explored the region in the immediate vicinity of the two marks. Precision varied from *S* to *S*, but with occasional exceptions the point could be located within a few millimeters.

The same procedure was then repeated with the other eye. Then the distances from each postocular projection to the mesial point of the head were measured. Except in a few anomalous cases, the *Ss* were examined only once.

The *Ss'* handedness and eyedness were also determined. If an *S* wrote, threw a ball, threaded a needle, or pushed a mop with his left hand, he was considered left-handed. Each *S* was given a slender cylinder to look through. According to the eye he used in sighting, he was classified as right- or left-eyed.

Subjects: (a) Aviation cadets. The postocular lines of regard were first obtained from 88 naval aviation-cadets. All had completed at least two years of college, were between 20 and 26 yr. of age and had excellent vision.

(b) Identical twins. To determine the role of hereditary components in the postocular lines of regard 30 pairs of identical twins were also examined. Four sets of twins were excluded from further study. One set had strabismus; one twin in another set had lost his left eye; the members of the third set had such poor vision that they could not see the straight edge; and the fourth set was excluded because one of the twins had multiple postocular lines. Of the 26 twinships included in the study 9 were monochorial and diagnosed at birth as identical by the physician who delivered them. The remaining 17 sets appeared to be and had always considered themselves identical. Blood tests, however, were not made; hence, the possibility exists that some of the sets may have been fraternal.

Four sets of twins wore glasses. With their glasses, they were able to see the straight edge clearly and their results were included. Fifteen of the 26 sets were women and 11 men. Ages ranged from 12 to 57 yr. All but one of the sets were white.

Results: (a) Cadets. The results for left and right lines, their midpoint

and the spread between them are presented in Table I. On the average, the two postocular lines are much closer to the midline than they are to their respective optical axes. In general, therefore, the somesthetic frame of reference dominates the visual in this situation.

Neither hand-dominance, eye-dominance, nor the interaction between them has any significant effect on the means in any of the four measures. Nor is there any heterogeneity of variance among the hand-eye dominance groups in any of the four measures. For the left line, however, the pooled variance of the two concordant groups, *i.e.* right-handed/right-eyed and left-handed/left-eyed, is significantly greater just

TABLE I

LEFT AND RIGHT POSTOCULAR LINES, THEIR MIDPOINT AND THE SPREAD BETWEEN THEM AS FUNCTIONS OF HAND- AND EYE-DOMINANCE

The units for means and variance-estimates are millimeters and square millimeters respectively

Group			Left line		Right line		Midpoint*		Spread	
handed	eyed	N	$\bar{X}$	S^2	$\bar{X}$	S^2	$\bar{X}$	S^2	$\bar{X}$	S^2
R	R	43	11.7	127.9	6.7	62.2	2.5*	27.4	18.3	273.5
R	L	21	10.8	68.0	3.5	62.9	3.7	22.8	14.3	170.4
L	R	10	11.1	102.1	6.6	94.0	2.3	29.0	17.7	276.2
L	L	14	14.8	149.9	6.3	94.2	4.3	21.2	21.2	403.8
Total		88	11.9	116.1	5.9	72.6	3.0	26.1	17.7	274.4

* The mid-point was calculated by subtracting the right from the left postocular line and dividing by two. Therefore, since all means for the mid-points are positive, all lie to the left of the midline.

beyond the 0.05 level than the pooled variance of the two discordant groups.

In all four dominance groups the mid-point of the two postocular lines lies to the *left* of the midline of the head. For all the naval cadets (88 *S*s) the average mid-point is 3 mm. to the left. This result is significant, against the hypothesis that the mid-point falls on the midline, well beyond the 0.001 level.

This finding was most pronounced. In 62 of the 88 cadets the mid-point fell to the left of the mesial line, sometimes by a considerable distance. In one cadet the left line lay 45 mm. to the left, while the right line fell on the midline; in other examples, the lines lay 30 and 5, 31 and 6, 23 and 0, and 27 and 5 mm. to the left and right sides respectively.

In 17 of the 62 cases *both* postocular lines lay to the left of the midline, in some instances by as much as 7 or 8 mm. In one instance, the left line lay 25 mm. and the right line 8 mm. to the left of the midline.

In 14 cases the deviations of the left and right lines to their respective sides were equal. In eight of these cases both lines fell directly on the midline.

In 12 cases the mid-point of the two lines lay to the right of the midline. In only two of these cases was the departure from the midline really pronounced. One cadet located the left line on the midline and the right line 13 mm. to the right; another cadet located the left line 4 mm. to the left and the right line 23 mm. to the right. This second cadet had heterochromia iridis. His right eye was hazel and his left eye brown. Nevertheless, to judge from the tests he had passed, his vision was perfectly normal. He was left-handed and right-eyed. Because of his unusual performance, this cadet was retested four days after the original testing.

TABLE II

LEFT AND RIGHT POSTOCULAR LINES, THEIR MIDPOINT AND THE SPREAD BETWEEN THEM

$N=26$ pairs of identical twins

Statistics	Left line	Right line	Midpoint	Spread
mean	8.8	4.8	2.0	13.5
variance (total)	80.3	58.1	18.0	204.8
variance (twins)	50.4	20.0	13.5	100.0
intraclass V	0.37	0.66	0.25	0.51

The results were very much the same. At retest, he located the left line 4 mm. to the left and the right line 26 mm. to the right.

In only two cases did a cadet locate both postocular lines to the right of the mesial line. In both instances the two lines fell together at 2 mm. to the right of the midline.

A final result is that the variance of the left postocular line was greater than the variance of the right ($P < 0.02$, two tails).[2]

(b) Twins. The results from the twins are given in Table II. Again, the average midpoint lies decidedly to the left of the mesial line of the head; significance reaches beyond the 0.01 level.

Among the twins spread between the two lines was less, though not significantly so, than in the cadet sample. Spread in the 4 twinships who wore glasses was half that in the remaining 22; again, however, the result was not significant. The spread among the twins was slightly less between the men than between the women. Age was quite unrelated.

In the bottom row of Table II, the intra-class correlations for the four measures are presented. The correlation for the right postocular line is

[2] The test used was for the difference between two correlated variances, Quinn McNemar, *Psychological Statistics*, 1955, 244.

significant beyond the 0.01 level. It is significantly greater at the 0.05 level than the correlation for the left line.[8] The correlation for the left line is significantly greater than zero at the 0.05 level.

(c) Anomalies. In one cadet and in one twin, several distinct postocular lines were observed for a single eye. This condition, which was more extensively studied in the cadet, was distinguished by three remarkable features. First, this *S* localized a touch on the *side* of the head as on the line. Secondly, with his right eye, for example, he localized at least two and sometimes three distinct lines. Third, he localized points well spread on *both* sides of the head as on the line.

The cadet responded immediately to a touch on the right side of his head as on the line; he was using his right eye. *E* continued to move over the top of his head toward *S*'s left. For the first two points the cadet said that the touch was to the left of the line. Then, as *E* continued to move toward the midline, the cadet said that the touch was to the *right* of the line. He then localized a second line on the midline. *E* continued to move leftward and again *S* reported that the touch was to the right of the postocular line. He localized a third line far out on the left side of his head. This procedure was repeated many times with essentially the same result. Sometimes the points would be somewhat differently localized and sometimes *S* would miss one of them, but always there were at least two lines and always they were well spread on both sides of the head.

Midway in these proceedings *E* noticed that the cadet was rubbing his eyes. When asked about it, he said that he thought he was "seeing things." He explained that it seemed to him that the ruler moved, *i.e.* that he saw it move without loss of alignment. When this first happened, he closed his eyes and looked again, but it continued to happen.

This cadet was examined on several occasions and his responses discussed at length with him. He soon realized the irrationality of what he was reporting and in the later tests stated that he knew better. If, however, I wished him to say how it seemed to him, then there was more than one line and the straight edge seemed to move. The cadet had no difficulty at all telling whether a touch on the skin was to his right or left; nor did he confuse two points homologously placed to either side of his backbone. It was only with respect to the visual straight edge that he became confused.

The cadet reported that he was originally left-handed and was trained to be right-handed. He still played pool, pushed, shot basketball, and ate with his left hand, but he wrote (miserably), threw a ball, and threaded a needle with his right hand. At six years of age he was going down hill on a sled, hit a tree, and suffered his first concussion. In the 8th grade he bashed into a wall playing basketball and suffered a second concussion.

The twin behaved very much the same way on examination, except that the lines on the contralateral side of the head were less pronounced. She said she had always been right-handed and right-eyed and reported no accidents involving her head. Her twin sister had uncomplicated single lines, well localized near the midline.

[8] The test used was for the difference between two correlation coefficients based on the same subjects, P. O. Johnson, *Statistical Methods in Research*, 1949, 54.

Several other *S*s behaved in a manner which resembled these two *S*s; but it was always transitory. On the first passage they might report a touch 20 mm. to the right of the midline as on the line. Later they would localize the line more definitely on the midline or near it. When the process was repeated, the "second line," *i.e.* the one off the midline, would not be reported. When the scalp was touched 20 mm. to the right, *S* would say that it was definitely to the right of the postocular line. Occasionally an *S* would report a touch on the side of the head as on the line. Usually he had failed to understand the directions. In no case, other than the cadet and twin with multiple lines, could the report be repeated; it no longer seemed that a touch on the side of the head was lined up with the straight edge.

In some *S*s there was an area of uncertainty that might be as broad as 10 mm. Any touch in this area seemed on the line. Some of these *S*s gave the impression that there might be two lines so close together that they fused in some way. In no case, however, again excepting the one cadet and one twin, were definitely distinct repeatable lines observed.

Discussion. The principal findings of this study are: (1) a general tendency for the midpoint of the two postocular lines to lie left of the midline of the head; (2) greater variance in the left postocular line; and (3) a heavier genetic contribution in proportion to the total variance in the right postocular line. These three results submit to a straightforward interpretation. Both postocular lines are affected by heredity and environment; the experiential component exercises its influence in moving both lines predominantly to the left; and this influence works more powerfully on the left postocular line than on the right. From these premises it follows that the left postocular line would be more variable than the right. Since it is the left line which is primarily affected, the variability associated with the effect is correspondingly greater. The scale of the effect is greater with respect both to means and variances in the left line. To the extent, however, that environmental effects inflate the variance of the left postocular line they reduce the relative magnitude of its genetic components; hence, the lower intra-class coefficients for the left line. Hereditary components in the left postocular line are diluted with more environmental variance than they are in the right line.

Summary

A straight edge was set up directly in front of, at the same level as, and pointing directly at *S*'s one open eye in monocular vision. *S* was then asked to indicate when a touch on his scalp seemed to him to be lined up with the straight edge. So defined, the right and left postocular lines were found generally to lie closer to the midline of the head than to the respective optical axes. The midpoint of the two lines lay consistently

to the left of the midline. The variance of the left was greater than the variance of the right postocular line. And intra-class correlation in a collection of 26 sets of identical twins was greater for the right line than for the left. Hand- and eye-dominance seemed not to bear upon the results. The interpretation was offered that both postocular lines are influenced by both heredity and environment, that environmental influences produce the leftward tendency of the two lines, and that this tendency is more powerfully felt in the left line.

DETECTION IN A HOMOGENEOUS VISUAL FIELD UNDER A CONDITION OF INFINITE DEPTH OF FOCUS

By JOHN CATALANO and MILTON S. KATZ
U. S. Naval Training Device Center

The ability of an observer to detect targets is relatively poor in a homogeneous visual field, *i.e.* a field lacking focusable structure such as might be encountered in space flight. Whiteside maintains that the decrement in detection is due to the fact that the *O*'s eye is accommodated at a distance of one to two meters and cannot voluntarily relax this accommodative state.[1] *O* is therefore myopic for targets at greater distances. The physical characteristics of targets which determine their detectability, such as size and brightness, must be greater than would be necessary if *O* were accommodated for the actual distance of the targets. Consequently, if *O*'s eye could be made to have infinite depth of focus, targets appearing at any distance should be easily detectable without requiring *O* to be accommodated for target-distance.

The depth of focus of the eye can be made effectively infinite with the use of an artificial pupil having a diameter somewhat less than the smallest natural pupil.

The purpose of this experiment was to study the effect of approximating infinite depth of focus, by use of an artificial pupil, upon target-detection in a homogeneous visual field.

METHOD

Observers. The *O*s in this study were three emmetropic men aged 27, 28, and 38 yr. who were experienced *O*s in studies on detection.

Apparatus. The homogeneous visual field was a white matte, wide-angle projection-screen which has been described in detail.[2]

The luminance of the field was 107.6 m.L. with a deviation over the field of 1.1 m.L. as measured with a Spectra Pritchard Photometer.

* Received for publication March 22, 1965. This work was performed as a part of the Human Factors Psychological Laboratory In-House Research Program, U. S. Naval Training Device Center, Port Washington, New York. The authors are indebted to Mr. Paul A. Cirincione for his cöoperation in this experiment.

[1] Thomas Whiteside, *The Problems of Vision in Flight at High Altitude,* 1957, 85-92.

[2] M. S. Katz, William Metlay and P. A. Cirincione, Effects of stimulus and field size on the accuracy of orientation in the homogeneous environment, *Percep. mot. Skills,* 20, 1965, 167-172.

O was seated in a *chaise longue* on a line bisecting the field. His eyes were 12 ft. 9 in. from the area on which the targets were projected within the *O*'s normal line of regard. The targets were circular spots of light which subtended visual angles of 3′ 36″, 4′ 12″, 4′ 48″, 5′ 24″, 7′ 12″, and 8′ 24″. These were projected on the screen in positive contrast by a Revere Type P-808 35 mm. slide-projector mounted on a tripod 5 ft. 6 in. above the floor. The distance from the projector to the screen was 16 ft. 4 in. A 0.2 D Wratten neutral density-filter was placed before the projector, resulting in 12.9 m.L. target-luminance.

Artificial pupils were made by drilling holes of 3 mm. and 1 mm. in aluminum sheeting which was .015 in. thick. Willson rubber, over-all goggles, Type X44, were modified to bring the pupils as close to the eye as possible. The closest distance that could be comfortably obtained was about 4 mm. from the eye. Field-brightness for the different size pupils was equated by placing a 0.6 D Wratten neutral density filter over the 3-mm. pupil while a 0.0 D Wratten neutral density filter was placed over the 1-mm. pupil.

The angular field of view was 20° 41′ 24″ for the 1-mm. pupil and 27° 3′ for the 3-mm. pupil. The *O*s could not identify which pupil they were looking through on the basis of any noticeable difference in brightness or extent of field.

Procedure. The experiment was designed to compare frequency of detection for the two artificial pupils. The 1-mm. pupil was used to approximate infinite depth of focus. Prior to the experiment, it was determined that infinite depth of focus could be approximated with a 1-mm. pupil. The criterion used in this determination was that an object at a distance 20 ft. and an object at a distance of 2 in. from the eye both remain in focus regardless of which object is fixated. It was therefore certain that any target appearing on the screen would be in focus regardless of the accommodative state of the eye. The 3-mm. artificial pupil served to control for any detrimental effect which the wearing of the goggles might have had upon *O* since depth of focus was not infinite with this pupil-size. Monocular detection was studied for the reason that it is more difficult to overcome he myopic effect when one eye is used than when both eyes are used.[8] Since the purpose of this experiment was to study the use of an artificial pupil in overcoming the myopic effect resulting from an involuntary state of accommodation, it was essential that the influence of voluntary relaxation of accommodation be minimized. The left lens of the goggles was so masked for all *S*s that no light was admitted to the left eye, which however could remain open.

An experimental session consisted of 88 trials, 44 trials with one artificial pupil size, a minimum of 3-min. rest and readaptation-period, and 44 trials for the second artificial pupil size. The order of pupil-conditions was counterbalanced for each *O*. Each experimental session was preceded by a familiarization-period during which the experimental targets were presented with 26.9 m.L. background-luminance and were easily detectable. This was done to familiarize *O* with the target-sizes that were to be presented later under reduced contrast ratio conditions. After the familiarization-period and preceding each pupil-condition, background-luminance was increased to 107.6 m.L. and *O* allowed 5 min. to adapt. During this time, a fixation-target of 57′ 36″ of arc was projected on the screen covering the area at which the experimental targets would appear. *O* so adjusted his head- and body-position that he was

[8] Whiteside, *op. cit.*, 76.

fixating the projection-area of the screen through the center of the artificial pupil. He was instructed to keep his head and body steady throughout the experimental period, and to report immediately any movement that would result in his gazing at an area of the screen other than one at which targets were to appear. In the few instances in which this occurred, the trials were repeated. Correct positioning was checked every 21 trials by presenting the fixation-target again. Each 21 trials consisted of 3 blocks within which each of the 7 targets was presented. For the first half of the experiment, targets were presented in a descending-descending-ascending order within each of the three blocks; and for the second half of the experiment in an ascending-ascending-descending order. Each target was presented automatically for 6 sec. with an intertrial-interval of 2 sec. within the region seen through the central portion of the pupil, although the exact location varied slightly from target to target. This variability imposed a very slight degree of search upon the detection-task, but since the target was within the area seen through the artificial pupil, only eye-movements were required and not head-movements.

Each target was presented six times per session for each size of the artificial pupil.

TABLE I

PERCENTAGE TARGET DETECTION FOR EVERY *O*

S	Target size (arc)													
	3′0″		3′36″		4′12″		4′48″		5′24″		7′12″		8′24″	
	1 mm.	3 mm.	1 mm.	3 mm.	1 mm.	3 mm.	1 mm.	3 mm.	1 mm.	3 mm.	1 mm.	3 mm.	1 mm.	3 mm.
MSK	0	0	0	4.1	16.6	41.6	83.3	87.5	95.8	87.5	91.6	91.6	100.0	100.0
PAC	0	8.3	4.1	33.3	12.5	41.6	70.8	75.0	70.8	79.1	91.6	79.1	95.8	100.0
JC	0	0	8.3	4.1	54.1	45.8	87.5	83.3	91.6	87.5	87.5	100.0	95.8	100.0
X	0	2.7	4.2	13.9	27.8	43.0	80.6	81.9	86.1	84.7	90.3	90.3	97.2	100.0

Every *O* was tested in four experimental sessions yielding a total of 24 replications per target for each artificial pupil size.

O was instructed to indicate verbally when he detected a target. Silence was recorded as non-detection.

RESULTS

Frequencies of target-detection for each *O* are shown in Table I. The mean for all *O*s plotted on normal deviate probit scale is shown in Fig. 1. As may be seen, detection was not enhanced with the use of the 1-mm. artificial pupil which approximated infinite depth of focus. If anything, detection was impaired under this condition for targets smaller than 4′ 48″ of arc for *O*s *PAC* and *MSK*. Detection was essentially similar under both pupil-conditions for *O JC*.

DISCUSSION

The hypothesis that the emtpy field, myopic effect could be overcome by inducing infinite depth of focus was not supported by the data. There are several possible reasons why these results were obtained.

(1) Poor detection and myopia may not be causally related in this case, but may be results of exposure to a structureless environment. Miller and Ludvigh have suggested that poor detection-performance in homogeneous visual fields may be due to disruption of normal search-patterns and lack of information as to location of the eye and areas previously searched resulting from the absence of structure.[4] If this is in fact the case, then

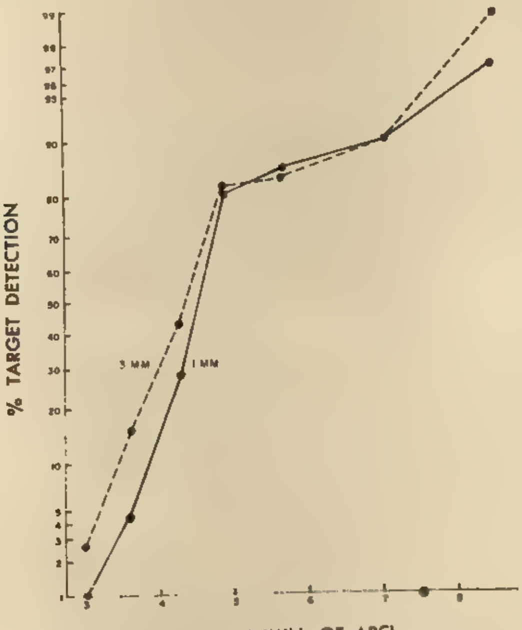

FIG. 1. MEAN PERCENTAGE OF THE TRIALS IN WHICH THE TARGET WAS DETECTED

providing infinite depth of focus would not be expected to result in enhancement of detection. On the other hand, the limitation of the field of view resulting from the use of artificial pupils *per se* might be expected to enhance detection because there is less area to be searched. The fact that no difference in detection was found regardless of the size of the pupil raises questions as to the adequacy of both views. Comparison of the data of the present study with those of an earlier one carried out in the same empty field indicates that detection for targets larger than 5 min.

[4] J. W. Miller and E. J. Ludvigh, Visual detection in a uniformly luminous field, *J. Aviat. Med.*, 29, 1958, 603-608.

of arc in size is comparable whether or not artificial pupils are used.[5] For targets smaller than 5 min. of arc in size, detection with the artificial pupils was poorer than had been found in the earlier study. The loss of acuity for the smallest targets may have been due to a combination of apparatus-factors including the fact that monocular viewing was used with the artificial pupils.

(2) It is possible that the artificial pupils in this experiment served to work against the hypothesis being tested. All the *O*s were aware of concentric rings and haze appearing at the edges of the artificial pupils. This is a phenomenon that is unavoidable when artificial pupils are used. Although the center of the pupils were clear and free of such interference, a greater proportion of the total field of view for the 1-mm. pupil than for the 3-mm. pupil lacked clarity. It was therefore more likely, should slight head-movement occur, that the target be occluded when the smaller pupil was used. A greater tolerance of head-movement was possible with the larger pupil. The plausibility of this explanation is reduced but not refuted by the fact that the *O*s were trained and aware of such a possibility and would without hesitation report any head-movement that might have occurred. It was not considered difficult to fix the head to insure that the targets would appear within the effectively clear area for the series of 21 trials before re-positioning occurred. Since no great difference in detection was found in spite of the factors ostensibly favoring the large pupil, the procedure was considered to have been valid.

(3) A possible reason for lack of superior detection with infinite depth of focus is that sufficient time was not allowed for the myopic accommodative state to occur with the larger pupil. Whiteside found that full, involuntary accommodation occurred within 60 to 80 sec.[6] Perhaps the use of the fixation-target and the 2-sec. intertrial-intervals was of too short duration to allow for involuntary accommodation to occur. The use of ascending and descending presentation of stimuli was decided upon to take this possibility into consideration. When the descending-descending-ascending order of targets-by-blocks presentation was employed, there was a period sometimes as long as 60 sec. between target-detections. This should have been sufficiently long for involuntary accommodation to occur; yet the 1-mm. pupil did not enhance detection any more than when there was a 2-sec. interval.

---

[5] Katz, Cirincione, and Metlay, Empty visual field studies: Some effects of corrective lenses, filters and structure, *NAVTRADEVCEN Technical Report IH-14*, 1964, 11-15.

[6] Whiteside, *op. cit.*, 93-97.

Conclusion

Three possible reasons have been offered to account for the fact that detection was not superior under the condition of infinite depth of focus. A logical extension of the hypothesis of Whiteside led to the expectation that the nullification of the effects of empty field myopia would have enhanced target-detection. This experiment has shown, however, that providing infinite depth of focus by the means employed does not result in the improvement of target-detection in a homogeneous environment.

Summary

It has been proposed by Whiteside that, in a homogeneous visual environment, involuntary accommodation results in a myopic condition which impairs target-detection. A means of overcoming this myopia by approximating a condition of infinite depth of focus with the use of an artificial pupil was studied. No improvement of target-detection resulted from this procedure.

TWO-CHOICE LEARNING: A FUNCTION OF INDICATED RANDOMNESS AND AMOUNT OF REWARD

By Sherman L. Guth and Raymond W. Frankmann
University of Illinois

It has been suggested that *Ss'* predictions in two-choice probability-learning may reflect a succession of hypotheses about the sequential structure of the series of events to be predicted.[1] As Edwards has noted, random series preclude regular confirmation of any hypothesis regarding the construction of the series. Therefore, *Ss* who are unaware that the sequence is random may seldom be expected to adopt the optimal strategy of always predicting the more frequent event.[2] Similar behavior might not be expected on the part of *Ss* who are convinced that the sequence is random.

Attempts to modify *Ss'* apprehensions of the structure of the sequence by instructing them that the sequence is randomly generated have produced typical probability-matching rather than trends toward optimal behavior.[3]

Rubenstein found that *Ss* required to predict whether cards drawn from an unseen deck would be blue or yellow exhibited typical learning curves and had proportions of terminal responses, which closely matched the 67:33 proportion of the two colors of cards. In contrast, *Ss* who predicted whether a king or a jack would be drawn from a set of three cards known to consist of two kings and a jack quickly adopted strategies which produced predictions of "king" on over 75% of the trials. In this case, *Ss* saw the cards being shuffled prior to each trial.[4] In a similar experiment, Nies found probability-matching for *Ss* who predicted which of two colors *E* would announce, more nearly optimal behavior for *Ss* led to believe they were predicting the color of marbles shaken from a box, and rapid approximation to optimal strategy for *Ss* who were told that the box contained 70 marbles of one color and 30 of another.[5]

---

* Accepted for publication October 30, 1964. This paper is based on the Master's thesis of Sherman L. Guth and was supported by grants from the National Science Foundation (G-16302) and the University Research Board of the University of Illinois.

[1] J. J. Goodnow, Determinants of choice-distributions in two-choice situations, this Journal, 68, 1955, 106-116.

[2] Ward Edwards, Probability learning in 1000 trials, *J. exp. Psychol.*, 62, 1961, 385-394.

[3] N. H. Anderson and D. A. Grant, A test of a statistical learning theory model for two-choice behavior with double stimulus events, *ibid.*, 54, 1957, 305-317; Howard Brand, J. M. Sakoda, and P. J. Woods, Effects of random versus pattern instructional set in a contingent partial reinforcement situation, *Psychol. Rep.*, 3, 1957, 473-479; E. D. Neimark and E. H. Shuford, *J. exp. Psychol.*, 57, 1959, 294-298.

[4] Irvin Rubenstein, Some factors in probability matching, *ibid.*, 57, 1959, 413-416.

[5] R. C. Nies, Effects of probable outcome information on two-choice learning, *ibid.*, 64, 1962, 430-433.

Morse and Rundquist obtained nearly optimal behavior from *Ss* who predicted whether a needle would intersect parallel lines ruled on a target upon which the needle was dropped (Buffon's needle-problem). Probability-matching obtained when the event-sequences generated in this way controlled which of two lights would occur on each trial.[6]

It is not clear, from the experiments discussed above, whether the instances in which nearly optimal strategies were observed resulted from *Ss*' understanding that the sequence was randomly generated or because of differences in the task between experimental and control conditions. In the present study, *Ss* were required to predict whether each of a series of cards, turned from the top of a large deck, would be marked *X* or *O*. Some groups were allowed to shuffle the deck prior to the series of predictions. Other groups had the same task, but were not allowed to shuffle the deck.

Since monetary rewards increase the tendency to use optimum strategies,[7] amount of reward was also varied in view of the possibility that an indication of randomness might not be effective in the absence of reward for correct predictions. A group which received a poker chip, of no monetary value, for each correct prediction was included to check on the possibility that accumulation of tokens, regardless of their value, might modify strategies by providing a visible record of successes.[8]

Method. Two levels of indicated randomness (shuffle and nonshuffle) were combined factorially with four amounts of reward (none, chip, 1¢, and 5¢) for correct predictions. In addition, half the *Ss* in each of these groups received *X* and half received *O* as the more frequent symbol. For each *S*, there were 180 trials, with an expected 75:25 ratio between the frequencies of the symbols.

Subjects. The *Ss*, 128 students from an introductory psychology course, were randomly assigned in equal numbers to the 16 experimental conditions.

Procedure and apparatus. The experiment took place in a small room with *S* and *E* seated opposite each other at a desk-sized table. The deck of symbol-cards was visible to *E* and *S*, but a screen prevented *S* from observing *E*. The *S* made his predictions immediately following the sound of a buzzer which occurred every 3.5 sec.

The *S* was required to turn over a card after each prediction and thus obtained immediate information as to the correctness of the prediction. Reward, when re-

[6] E. B. Morse and W. N. Runquist, Probability-matching with an unscheduled random sequence, this JOURNAL, 73, 1960, 603-607.

[7] Edwards, Reward probability, amount, and information as determiners of sequential two-alternative decisions, *J. exp. Psychol.*, 52, 1956, 177-188; Goodnow, *op. cit.*, 116; Sidney Siegel and D. A. Goldstein, Decision-making behavior in a two-choice uncertan outcome situation, *J. exp. Psychol.*, 57, 1959, 37-42.

[8] J. L. Myers, R. E. Reilly, and H. A. Taub, Differential cost, gain, and relative frequency in a sequential choice situation, *ibid.*, 62, 1961, 357-360.

quired, was accomplished by sliding a chip, a penny, or a nickel to *S*'s side of the table.

Symbol-cards were made from blank playing cards. An *X* or an *O*, each about ½ in. high and ¼ in. wide, was lightly stamped in ink on one side of each blank. The symbol was not detectable on the reverse side of the card.

To minimize end-effects, the decks presented to *S* consisted of 220 cards. Between experimental sessions, *E* repeatedly cut and shuffled each deck.

After each experimental session, *S*s were urged not to discuss any aspect of the experiment with acquaintances. They were also asked if they had previously heard anything about the experiment; *S*s in the monetary reward groups were assured that they could keep their winnings, even though they did have prior knowledge of the experiment. (Only one *S* gave an affirmative answer to the question, and her data were discarded.) In addition, the *S*s were asked to discuss the strategies they had used in the experiment.

Instructions. In the instructions, which were read aloud by *E*, it was stated that that experiment concerned decision-making, in general, and the prediction of uncertain events, in particular. Each *S* was told to make as may correct predictions as possible.

During the explanation of the task, instructions appropriate to the various groups were inserted with minimal changes of wording. The *S*s in the shuffle-groups were told that the cards were arranged randomly and were instructed to shuffle the cards to assure themselves that no pattern existed in the series. No remarks about the nature of the series were made to *S*s in the nonshuffle-groups.

The *S*s in the monetary reward groups were told how much they would win for each correct prediction. A concerted effort was made to convince the *S*s that they could expect to win a substantial amount and that they would keep their winnings (which they did). The *S*s in the chip-groups were told that they would return the chips to *E* after the experiment.

Results. During the first 20 trials, the proportion of predictions of the more frequent symbol was between 0.53 and 0.59 for the eight main groups. Proportions increased over successive blocks of trials, but the monetary reward groups did not reach asymptotic levels of performance. The proportions for the nonmonetary groups appeared to be stable after 120 trials and the proportion of predictions of the more frequent symbol made by each *S* during the last 60 trials (this proportion will be denoted *PMF*) was used as the main dependent variable. Values of *PMF* are shown in Table I. Since no significant effects were associated with whether the more frequent symbol was *X* or *O*, the data for these two conditions have been combined.

Significant effects in an analysis of variance of the *PMF*s were all associated with the amount of reward variable ($F(3,112) = 3.97$, $p < 0.05$).[9] More interesting details were revealed by orthogonal comparisons

[9] Within-group variances ranged from 0.0072 to 0.0084 for the eight main groups and from 0.0068 to 0.0189 for the 16 groups of the complete design. In

among the means for the four levels of this variable.[10] Comparison of the monetary with the non-monetary groups yielded $F\ (1{,}112) = 8.48$ ($p < 0.005$); for chip *versus* no reward the value of F was 0.11 and for 1¢ *versus* 5¢ reward the value was 0.64. Other comparisons suggested by the data were not significant when tested by Scheffe's method.[11]

In the groups that received 5¢ rewards, six *S*s in the shuffle-group and two *S*s in the nonshuffle-group had *PMF*s greater than 0.90. Application of the Kolmogorov-Smirnov test, however, indicated no significant differences between the *PMF* distributions of the shuffle- and nonshuffle-groups, either for the various amounts of reward or for all groups combined.[12]

As a check on the method of generating sequences of symbols, it was noted that the proportion of the last 60 trials on which the more frequent symbol occurred ranged from 0.67 to 0.88 for individual *S*s and from

TABLE I

Mean Proportion of Predictions of the More Frequent Symbol during the Last 60 Trials

Indicated randomness	Reward: 0	chip	1¢	5¢	Means
shuffle	.73	.78	.82	.85	.79
nonshuffle	.78	.75	.79	.81	.78
Means	.76	.76	.80	.83	

0.74 to 0.77 for the eight main groups. The within-group correlation between these proportions and the *PMF*s was significant ($r = +.76$, $p < 0.01$) only for the nonshuffle-chip group; the nonsignificant positive and negative correlations for each of the seven other groups showed no systematic relation to either of the independent variables.

Discussion

That monetary reward increased the tendency to use optimum strategies is in accord with results of other relevant studies.[13] The similarity of the groups that received poker chips to the groups that received no reward, except knowledge of results, indicates that the effect of reward is not

neither case did Hartley's F_{max} statistic (B. J. Winer, *Statistical Principles in Experimental Design*, 1962, 93) indicate significant heterogeneity of variance. There was no indication of a systematic relation between means and variances.

[10] Winer, *op. cit.*, 65-70.

[11] Winer, *op. cit.*, 88-89.

[12] Siegel, *Nonparametric Statistics*, 1956, 127-131.

[13] Edwards, Reward probability, amount, and information as determiners of sequential two-alternative decisions, *J. exp. Psychol.*, 52, 1956, 177-188; Goodnow, *op. cit.*, 106-116; Siegel and Goldstein, *op. cit.*, 37-42.

due simply to the receipt and accumulation of tokens for correct predictions. The small difference between the effects of 1-cent and 5-cent rewards agrees with the findings of Meyers, *et al.* that *PMF* increases only slightly when reward is increased from 1 to 10¢.[14]

The absence of any effects due to indicated randomness supports the conclusion that, within a single task, *S*s who must develop strategies by observing the consequences of repeated predictions do not readily accept, or understand, information that the sequence is randomly generated. It is possible, of course, that effects due to this variable or an interaction with amount of reward woud have developed had there been more trials.[15]

The informal verbal statements given by *S*s after their participation in the experiment provide some clues as to why the randomness variable had no effects. As was expected, essentially all of the *S*s in the nonshuffle groups thought that some pattern existed and attempted to determine its nature throughout the 180 trials.

It was surprising, however, that many *S*s who personally shuffled the cards indicated that they were thinking in terms of a pattern. That is, they would say something like, "It seemed that there would be three *X*'s, then an *O* then four *X*'s, then an *O*, etc." Additional questioning elicited responses such as "I know the cards were mixed, but they seemed to be occurring like that [in a pattern] anyway."

The verbal reports of other *S*s in the shuffle-groups suggested that they knew that no pattern existed, but simply did not realize that the optimum strategy would be always to predict the more frequent symbol. Still others recognized the optimum strategy, but occasionally predicted the infrequent event "just for fun." One *S* remarked that he thought he would be cheating if he always guessed *X*.

---

[14] J. L. Myers, J. G. Fort, Leonard Katz, and M. M. Suydam, Differential monetary gains and losses and event probability in a two-choice situation, *ibid.*, 66, 1963, 521-522.

[15] Edwards, *op. cit.*, *J. exp. Psychol.*, 62, 1961, 385-394.

THE POTENCY OF A LIGHTNESS-ANCHOR AS A FUNCTION OF THE REFLECTANCE OF ITS BACKGROUND

By WILLIAM BEVAN and EDWARD D. TURNER, Kansas State University

The paradigm of the anchor-experiment has been an important vehicle in the development of frame-of-reference psychophysics. Helson early identified it as a convenient means of quantitatively assessing the role of context in the prediction of psychophysical judgments.[1] The extent of the contribution of an anchor to any representative judgmental value (*e.g.* the *I.P.*) may be expressed as the product of a constant identified with its function (context vs. focal stimulus), its physical magnitude, and the frequency of its occurrence.[2]

The usual anchor-experiment, however, shares with the experimental design of traditional psychophysics the methodological assumption that, for purposes of prediction, stimuli are unidimensional; and all dimensions, except the one being judged, are held to be constant or otherwise controlled. Under such circumstances, the physical magnitude of the anchor, relative to the magnitudes of the stimulus-series, constitutes a competent index of its potency. But a major value of frame-of-reference psychophysics has been its distinction between physical magnitudes and behaviorally effective (*e.g.* judged or 'apparent') magnitudes and its demonstration that these are only relatively infrequently the same value.

Accordingly, the purpose of this experiment has been twofold: (1) to add yet another demonstration of this important difference between simple physical definitions of stimuli and definitions functionally referenced to response; and (2) to show the dependence of one distinct class, the anchor-stimuli, upon apparent magnitude for its effectiveness.

The significance of background in the determination of effective magnitude has been recognized generally since the heyday of Gestalt psychology with its interest in the perceptual constancies. For example, Brown demonstrated that apparent velocity depends upon rate of physical displacement relative to the size of the

---

* Received for publication November 16, 1964. Supported by Contract Nonr-3634(01) between Kansas State University and the Physiological Psychology Branch, Office of Naval Research.

[1] Harry Helson, Adaptation-level as frame of reference for prediction of psychophysical data, this JOURNAL, 60, 1947, 1-29.

[2] Helson, *op. cit.*, 3; William Bevan and C. L. Darby, Patterns of experience and the constancy of an indifference point for perceived weight, this JOURNAL, 68, 1955, 575-584.

area within which displacement occurs.[3] More recently, Wallach has shown that perceived color depends upon contrast with the surround.[4] When a disk of light was projected on an otherwise dark screen, it appeared luminous but yielded no impression of grayness. When, however, an annulus of light was so projected as to surround the disk, the apparent brightness of the disk varied with the difference in intensity between disk and ring. Similarly, apparent size depends upon the relative size of figure and frame of reference over a wide range of projective sizes (Rock and Ebenholtz) and judged angularity upon the size of a prior series of produced angles (Jennings and Johnson).[5]

In recent years, studies of the relational determination of apparent magnitude have been marked by the quantitative specification of significant stimulus-context relationships. Thus Künnapas reports that apparent length of line varies linearly with the logarithm of the area of the frame of reference.[6] Adaptation-level theory, of course, provides a comprehensive quantitative treatment of the problem of apparent magnitude.[7] Apparent magnitude is expressed as the difference between the magnitude of a focal stimulus and an internal norm, in turn specified in terms of the interaction of three classes of variable including background.

The importance of the concept of effective magnitude is not limited to experiments on perception and judgment. Riley has reported that efficient transposition of a brightness-discrimination response by white rats is related to invariance in the relative brightness of the discriminanda and their ground.[8] Helson and Kaplan have also demonstrated the importance of effective magnitude for successful transposition.[9] Human *S*s were shown a series of five gray patches on a black background and were trained to identify a particular patch. They were then tested with a white background and 61% of the *S*s selected a patch that *looked* the same on the test-background as the 'correct' patch did on the black training background. These results were almost exactly reproduced by Weidenfeller, Hardesty, and Bevan.[10] Similarly, Bevan and Adamson have shown that human maze-learning relates to effective magnitude, in distinction to physical magnitude, of reinforcement.[11] Performance on a bolt-head maze was better for *S*s given pretest exposure to shocks weaker than the reinforcing shock when compared to that of *S*s given

---

[3] J. F. Brown, The visual perception of velocity, *Psychol. Forsch.*, 14, 1931, 199-232.

[4] Hans Wallach, Brightness constancy and the nature of achromatic colors, *J. exp. Psychol.*, 38, 1948, 310-324.

[5] Irvin Rock and Sheldon Ebenholtz, The relational determination of perceived size, *Psychol. Rev.*, 66, 1959, 387-401; J. W. Jennings and D. M. Johnson, Context effects in production and judgment, *J. Psychol.*, 56, 1965, 53-59.

[6] T. M. Künnapas, Influence of frame size on apparent length of line, *J. exp. Psychol.*, 50, 1955, 168-170.

[7] Helson, *Adaptation-Level Theory*, 1964, Ch. 4.

[8] D. A. Riley, The nature of the effective stimulus in animal discrimination learning: Transposition reconsidered, *Psychol. Rev.*, 65, 1958, 1-7.

[9] Harry Helson and Sylvan Kaplan, cited in Sigmund Koch (ed.), *Psychology: A Study of a Science*, 1959, 595-596.

[10] Linda Weidenfeller, Don Hardesty and William Bevan, The role of stimulus context in the performance of a lightness-discrimination task, *J. gen. Psychol.*, 72, 1965, 229-232.

[11] William Bevan and Robert Adamson, Reinforcers and reinforcement: their relationship to maze performance, *J. exp. Psychol.*, 59, 1960, 226-232.

pre-test shocks stronger than the reinforcer. The same relationships have been shown to obtain for running speed in a straightaway.[12]

Engel and Parducci have investigated the effect of variation in the background of some members of a psychophysical series upon judgments of other members of the series.[13] In an experiment in the categorical judgment of size, a change in the size of the background of one of the series members produced an inverse shift in the adaptation-level.

The present experiment examined the influence of background upon the potency of an anchor-stimulus. This relationship was explored over the full range of possible backgrounds with a view toward identifying the nature of the function that best describes it.

Method. (1) Design. The dimension chosen for judgment was lightness (brightness). The variable selected for experimental manipulation was the reflectance of the anchor-surround. The anchor-stimulus selected was a dark gray patch lying below the lightness-range of the series-patches. The experimental prediction was that when the anchor-surround had a reflectance higher than that of the stimulus-series, the influence of the anchor upon the judgments in the series would be enhanced; conversely, when the reflectance of the anchor-surround was below the level of the series-background, the influence of the anchor would be reduced.

The design involved six independent groups. Group *C* (the control group) judged the lightness of a series of five gray patches, each on a medium gray background (39% reflectance). Group *St.*, representing the standard anchor-condition, received, in addition to the series, the dark gray anchor-patch. All stimuli were presented on the same medium gray ground as Group *C*. The remaining four groups (Groups *W*, *LG*, *DG*, and *B*) were replicates of *St*, except that the backgrounds for the anchor were respectively white (73% reflectance), light gray (54% reflectance), dark gray (12% reflectance) and black (1% reflectance).

Subjects. The *Ss* were 60 women, all undergraduate volunteers from the course in General Psychology. An equal number were assigned at random to each of the six groups. All were untrained in observation and unfamiliar with the purpose of the experiment.

Materials. The stimulus-patches were squares, 7.5 cm. on a side, mounted on cardboard slides and presented in one field of a Gerbrands tachistoscope. The series consisted of Munsell values, light to dark, of 7.5, 7.0, 6.5, 5.5 and 5.0. The dark gray anchor had a Munsell value of 3.5.

Procedure. Only one field of the tachistoscope was used. The exposure-time for each stimulus was 0.5 sec. The over-all level of illumination in the laboratory was dim. Each *S* was seated before the tachistoscope and told that she would be shown a series of gray squares and that she would be asked to say "light" or "dark," whichever she ,udged the squares to be. To help her express her judgments, she was presented an 11-point scale varying from + 5 (light) through 0 (medium) to

---

[12] Roger Black, Robert Adamson, and William Bevan, Runway behavior as a function of apparent intensity of shock, *J. comp. physiol. Psychol.*, 54, 1961, 270-274.

[13] Gloria Engel and Allen Parducci, Value of background in the specification of the stimulus for judgment, this JOURNAL, 74, 1961, 569-575.

— 5 (dark). She was instructed to give her judgments as "light 5," "dark 3," "medium zero," etc. She was told further she could add categories to either end of the judgmental scale as the judgmental task required.

S then received a familiarization-sequence of 14 trials—two randomized blocks of 7 presentations each in which each series member appeared once and the anchor-stimulus twice. (The anchor was never formally identified for *S*. Its designation as such in the experimental design is based on its separation from the series on the intensive scale and its more frequent presentation.) Next, *S* was asked if she had any questions about the task; and when *E* was satisfied that *S* was competent to proceed, 5 additional blocks of 7 randomized trials each were presented as a continuous sequence. Individual trials were spaced at about 5-sec. intervals.

Results. (1) Basic anchor-effect. Each *S*'s judgments of lightness were summarized as the mean of all series members on all trials. This value for the control group (Group *C*) was 0.14, not deviating from the zero of the medium scale-value by more than a small and chance amount. Groups *C* and *St*, when taken together, constitute the traditional anchor-experiment. Results indicate the commonly-observed contrast-effect. Addition of the anchor-stimulus 1.5 Munsell categories below the series on two of every seven trials produced an average increase in apparent lightness of the series of 1.20 judgmental categories ($C - St. = 1.20$, $F_{1,18} = 9.75$, $P < 0.01$).

(2) Influence of the anchor-background. Fig. 1 presents the mean lightness-judgments for the several anchor-conditions. Group *St*, the standard anchor-condition, is represented by the third or middlemost point. The other points are values for Groups *B*, *DG*, *LG*, and *W*, respectively. It is thus manifestly clear that the power of the anchor is directly related to the lightness of its background and the hypothesis stated earlier is confirmed. Indeed, the use of the black background with the dark anchor reduced its potency almost to zero ($C - B = 0.21$ category units). Some interesting quantitative comparisons are in order. At the lower extreme of the plotted function, reflectance of the background has been reduced 38% and anchor-potency is only 17% as powerful as that produced under the standard anchor-condition. On the other hand, at the upper end, reflectance has been increased 34% and the anchor-effect is 60% greater than that for the standard condition. Between Groups is, of course, a highly significant source of variance ($F_{4,45} = 6.63$, $p < 0.001$) and the curve of best fit is a straight line. The linear component of the between-groups source yielded an *F* of 26.45 (with *df* of 1 and 45, $p < 0.001$) and accounted for 99.7% of the variance.

The implication of these data is clear. Anchor-potency in this situa-

tion is not simply a matter of the physical intensity of the anchor or of the distance by which it is separated from the series on the dimension being judged. Rather, it is a matter of its lightness relative to the light-

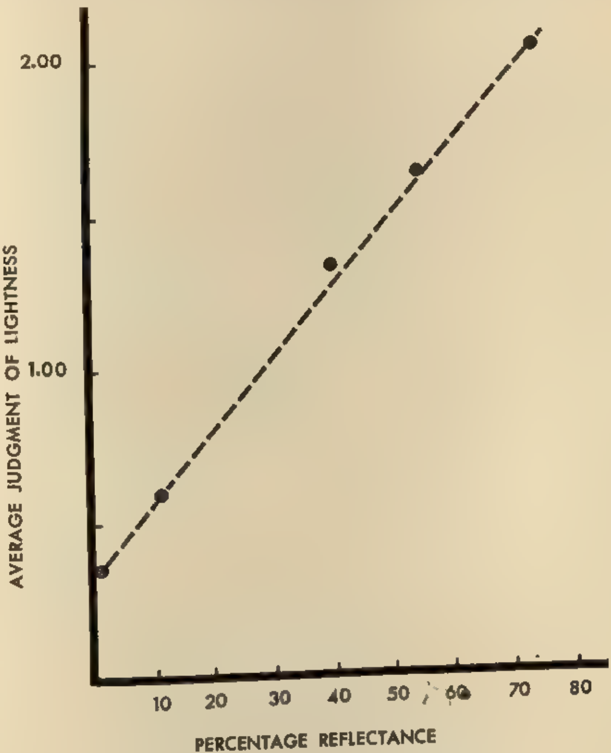

FIG. 1. ANCHOR-POTENCY AS A FUNCTION OF THE LIGHTNESS OF THE BACKGROUND

Data are mean lightness-judgments for an identical series when an anchor below the series appeared on a black, dark gray, medium gray, light gray or white background. Series members always appeared on the medium gray ground.

ness of its surround, and any general quantitative expression derived to predict the Adaptation-Level (*A-L*) for lightness must recognize this relationship. It may also be assumed to have its analog in other judgmental settings. It is interesting to note that none of the *S*s gave any indication of being aware of a difference in background-lightness between anchor and series.

SUMMARY

Sixty *S*s, divided at random into six groups of 10 each, judged the lightness of gray patches presented one at a time in a Gerbrands tachistoscope. One group, the control, received only the stimulus-series on a medium gray ground. The other five received, in addition, an anchor below the series in lightness, one on a background with reflectance equal to that of the series, the others each on a lighter or darker ground. Results indicate that the magnitude of the effect of the anchor was directly related to the relative lightness of the ground. The greatest effect of the anchor was produced when the dark gray anchor appeared on a white background; the weakest when the dark gray anchor appeared on a black background.

LIMITED MODIFICATION OF THE TRAPEZOIDAL ILLUSION WITH EXPERIENCE

By Ralph Norman Haber, University of Rochester

Ames devised the rotating trapezoidal illusion as a demonstration of the effect of the perceiver's assumptions about the visual world upon his perceptual experience.[1] The predominance of rectangular distal forms, Ames argued, causes perceivers (nearly always correctly) so to bias their phenomonal interpretations of proximal trapezoidal forms as to experience them as rectangular. This bias toward rectangularity augments a depth-effect in the perception of a trapezoid, the tall end being seen nearer to the perceiver than the short end. When the trapezoid is rotated, the perceiver, under this assumption or bias, has a consistent experience of oscillation; the tall end is always perceived as near and never is seen to go behind the short end.

There are few experimental data bearing on this crucial transactionalist principle—that assumptions about the properties of the distal forms determine the experience and not just the interpretation of the pattern of proximal stimulation.[2] Allport and Pettigrew, in one of the few direct tests, presented the trapezoidal illusion to Zulus, who were thought to live in an environment that would not be conducive to the formation of assumptions about rectangular distal objects.[3] Their results were equivocal, though Slack has suggested that perhaps even the Zulus did not lack rectangular assumptions.[4] If Slack is correct, then probably the theory cannot be tested in this manner with any group of human beings.

Kilpatrick reports anecdotally that the perceiver's knowledge about the various demonstrations, including the trapezoidal illusion, has little effect

---

* Received for publication December 26, 1964. This research was supported in part by grant MH-03244 from the Public Health Service. The work was done at Yale University. The author is indebted to Randall Bolsinger, Frederick Sakamoto, and John Gunnell, for assistance in the collection of data.

[1] Adelbert Ames, Jr., Visual perception and the rotating trapezoidal window, *Psychol. Monogr.*, 65, 1951 (No. 324), 1-31.

[2] W. H. Ittelson, Perception and transactional psychology, in Sigmund Koch (ed.) *Psychology: The Study of a Science*, 4, 1962, 660-704.

[3] G. W. Allport and T. F. Pettigrew, Cultural influences on the perception of movement: The trapezoid illusion among Zulus, *J. abnorm. soc. Psychol.*, 55, 1957, 104-120.

[4] C. W. Slack, Critique on the interpretation of cultural differences in the Ames trapezoid, this Journal, 72, 1959, 127-131.

on his perceptual experience.[5] Furthermore, psychologists who operate the apparatus generally report informally that their own illusions persist over many years. These are somewhat surprising findings in view of the implied experiential antecedents for the illusion. This experiment was designed, therefore, to determine whether the perceptual experience of the trapezoidal illusion can be modified by controlled experience and by disconfirmatory knowledge of any prior assumptions made by the perceiver.

Method. *S* was seated near one end of a long narrow room and viewed the trapezoid binocularly from a distance of 20 ft. The trapezoid was mounted at eye-level, 3 ft. from the other end of the room. Directly in front of it was a full partition, containing a 3×3 ft. guillotine window through which, when it was open, *S* could see the trapezoid. The figure, milled from $\frac{1}{32}$ in. aluminum, was $15\frac{7}{8}$ in. long; and $19\frac{1}{4}$ in. high at the tall end and $9\frac{1}{2}$ in. at the short end. Shadows were provided by black masking tape along the back edge of each vertical surface. The figure was illuminated indirectly by a 25-w. lamp placed overhead, thus eliminating all moving shadows. The back wall, visible when the window was open, appeared as a background (*Ganzfeld*) of indeterminate distance. During trials, there was no other illumination in the room. When the window was opened at the start of each trial, the trapezoid was always in the frontal plane relative to *S*, with the tall end to the left, rotating at 5 r.p.m. in a counter clockwise direction as viewed from above. This arrangement was suggested by Ames as ideal for the production of the illusion.[6]

All the *Ss* (73 in number) were Yale undergraduates, drawn from the introductory psychology course, and were tested individually and were remunerated for their services. When *S* was seated, he was read the following instructions.

When the large window in the back wall is opened, you will see a figure in motion. As quickly as you can, indicate the direction that the part nearest to you is moving, by calling out 'right' or 'left.' If it changes, then, as quickly as you can, call out the new direction.

For the illusion of oscillation to be complete, changes in direction would have to be reported every 6 sec. There were five experimental conditions:

(1) Control (N = 10). Seven 3-min. viewing trials were presented, separated by 1-min. rest-periods. No additional information was given. Calls of 'R' and 'L' were recorded, with the time in seconds.

(2) Gradual information (N = 28). Specific information was given after certain trials regarding the characteristics of the trapezoid and its motion. The information was printed on different pages of a booklet, and *S* was told at each point which page to read. After Trial 2, *S* read the correct thickness of the figure and that the shadows were painted. After Trial 3, he read that the shape was trapezoidal and was given its dimensions. After Trial 4, he read that the motion was always rotational to the right at 5 r.p.m. After Trial 5, *S* was permitted to examine the trapezoid and to resolve any doubt about its shape and size. After Trial 6, acting as if the experiment

[5] F. P. Kilpatrick, Assumptions and perception: Three experiments in Kilpatrick (ed.), *Human Behavior from the Transactional Point of View*, 1952.
[6] Ames, *op. cit.*, 17.

was over, *E* entered into a discussion with *S*, telling him the history of the illusion, the transactionalist theory, some of the experimental evidence, and the purposes of the present experiment. The *S* was encouraged to discuss his own impressions and performance, and to add his own ideas. Each *S* was questioned regarding any previous exposure to the illusion in classes or elsewhere. After this discussion, which lasted about 10 min., *S* was given another trial. Before being given the above listed items of information after Trials 2, 3, and 4, *S* was asked to describe how that particular aspect of the trapezoid about which he was to be informed appeared to him. He wrote his answers on preceding pages in the same booklet. After a period ranging from a week to a month, 15 *Ss* whose illusions had been modified returned for another seven trials. Prior to the first of these trials, there was a recapitulation of the discussion. Nine of these *Ss* were retested a third time from 10 min. to 6 hr. later.

(3) Early information (N = 10). *S* was engaged in the discussion and given the opportunity to handle the trapezoid before the first trial. Then there were seven trials without further information.

(4) Late information (N = 18). The first five trials were like those of the control condition. After the fifth trial, *S* was engaged in the discussion and, after the sixth, he was permitted to handle the trapezoid. For this group, the order of the fifth and sixth trials was reversed relative to that for the gradual group, and it differed from the gradual group also in that no information was given during the first five trials.

(5) Distance-variation (N = 7). After being given all the information and hearing discussion at the beginning, *S* was brought up to the opened window and, with the trapezoid rotating, he was encouraged to back up to his chair while trying to maintain the rotational experience. If there was a change to oscillation, he was then to come forward slowly, until he saw rotation, and then back up again. Distances from the partition were recorded.

In all conditions, *E* remained out of sight behind the partition, where he recorded *S*'s reports.

Results: (1) Control. Eight of the 10 *Ss* saw the illusion at full strength during each of the seven trials (30 oscillations per trial). No tendency for the illusion to weaken was noted, either in the number of oscillations or in their times of occurrence. The other two *Ss* had no illusion (*i.e.* made no reports) on the first trial or thereafter. Thus, 80% of the *Ss* in this condition were susceptible to the trapezoidal illusion. No change was found in the strength of the illusion over 21 min. of viewing.

(2) Gradual information. The percentage of *Ss* with the illusion at full-strength on the first trial was 89%, quite comparable to the Control *Ss*, but by the end of seven trials only 36% still had the illusion. (In general, the *Ss* either made 30 reports of change (oscillation) or none (rotation), though there were a number of cases in which there were 10-15 reports. Sometimes this 'intermediate' pattern would continue for several trials, but it never was followed by an increase in the number

of reports.) Since there was little reduction in the illusion until after the fifth and sixth trials, it seems that handling the figure and participating in a discussion of it were the most relevant experiences.

Fifteen of the 18 *Ss* who showed some breakdown in the illusion were retested. Of these, 11 continued to have vertical perception. Of the other four, one lost the illusion again after three additional trials. Nine of the 12 *Ss* without the illusion at the end of the second session were retested again, and none showed any vestiges of the illusion. Thus, once the illusion was destroyed, rotational experience was quite stable.

(3) Early information. When *Ss* were informed at the beginning, only 40% showed the illusion. This percentage agrees closely with that on the last trial of the gradually informed group, suggesting that the manner in which the information is given is not important.

(4) Late information. The first five trials, which were similar to those of the control group, yielded comparable data, although there was some weakening of the illusion during those trials (88-61%). Far sharper drops occurred after the fifth (discussion) trial (61-28%) and after the sixth (handling) trial (28-11%). Comparable trials for the gradual-information condition showed a drop from 68-50% after handling (Trial 5) and then from 50-36% after discussion (Trial 6). The two conditions were quite similar prior to the end of Trial 5, whch suggested that the earlier information was little better than mere viewing experience, but, giving information suddenly (as in the Trial 5) seemed more effective than having given some lesser information leading up to it—a total drop from 61-11% as compared with 68-36% ($p < 0.05$).

(5) Distance-variation. No *S* could move back the 20 ft. to his chair and still see rotation. The farthest any *S* could get was 16 ft., while the median distance was 8 ft.—less than half-way. In no instance did any *S* require a forward movement of more than 2 ft. to see rotation again, and in most cases it was less than 1 ft. The experience of oscillation occurred at about the same place as it had previously when *S* again backed up. Thus, trying to maintain the experience of rotation while backing up was ineffective; it did not aid at all in reducing the effects of the illusion, even if *S* had been given total information.

Less than half the *Ss*, distributed evenly over the conditions, reported having seen a demonstration of the illusion previously. No relationship was found between this knowledge and any outcome of this experiment. The written answers to the questions predicted a breakdown in the illusion only for the question on the nature of the motion of the figure; those *Ss* who said they saw it rotate did in fact make fewer reports

($X^2 = 20.6$, $p < 0.001$). Interestingly enough, while nearly half the *Ss* thought the figure was a trapezoid after the second trial, this knowledge did not facilitate a breakdown in the illusion; neither, apparently, did simply being told it was a trapezoid.

Discussion

While bearing directly on the validity of transactionalist theory, these results indicate that relevant information and exposure can affect the perceptual experience of the trapezoidal illusion. From 80-90% of the *Ss* were susceptible to the illusion in the beginning and nearly all of them remained so if given no specific information; but, when relevant information was provided, the illusion broke down for the majority of *Ss*. These results stand in contrast to the anecdotal reports of the extreme stability of the illusion in highly experienced and knowledgeable viewers. The basis of these anecdotal reports of stability remains to be determined.

Summary

Anecdotal reports in the literature suggest that the Ames trapezoidal illusion is not modifiable by extended experience with it or specific knowledge about it. Nevertheless, the *Ss* of this experiment gained veridical perception when systematically exposed to information about the illusion. The breakdown in the illusion persisted without further experience for periods as long as a month. The components of the knowledge were analyzed, as was the influence of the method of presenting it.

SATIATION AND THE TILT AFTER-EFFECT

W. B. TEMPLETON, I. P. HOWARD, and G. EASTING,
University of Durham, England

The problem of whether Gibson's spatial adaptational effects can be subsumed under Köhler's theory of satiation has received considerable attention in recent years. For example, Morant and Mistovitch, contrary to the prediction from satiational theory, have confirmed the occurrence of Gibson's 'indirect effect' on one spatial axis (apparent vertical or horizontal) of an inspection-line somewhat tilted from the other axis.[1] Also, Rich and Morant have shown that when the magnitude of the after-effect of tilt is plotted as a function of angle of inspection-figure, the resulting curve agrees more closely with the prediction that both types of effect are occurring than with the prediction from satiational theory alone.[2]

Two major points of conflict between the theories remain to be investigated: (1) The normalization of a single tilted line to the vertical or horizontal, which is the crux of Gibson's theory, cannot without additional assumptions be explained on satiational theory. Though frequently reported incidentally by *S*s in this type of experiment, normalization has never been rigorously demonstrated. (2) The second point, which is investigated in this paper, arises from the fact that whereas the occurrence of an after-effect of tilt requires that the inspection-line be tilted, a figural after-effect should occur between any pair of lines which form an angle with one another. Hence any effect of a vertical or horizontal inspection-line on a tilted line can only be a figural after-effect, whereas both theories predict an apparent repulsion of a vertical or horizontal line by a tilted inspection-line. If, therefore, there is a tilt after-effect as distinct from a figural after-effect, the two may combine in the second case to give a larger total effect than that obtained in the first case. If, on the other hand, there is no independent tilt after-effect, the effects in the two cases should be the same size, viz. that of the figural after-effect.

*Received for publication December 9, 1963. This research was supported by D.S.I.R. Grant No. 1208/69/1.

[1] R. B. Morant and Mildred Mistovitch, Tilt after-effects between the vertical and horizontal axes, *Percept. mot. Skills*, 10, 1960, 75-81.

[2] J. H. Rich and R. B. Morant, A two-factor approach to the study of tilt after-effects, Paper presented to the Eastern Psychological Association, 1960.

Köhler and Wallach carried out this experiment and found that with two lines, one vertical (or horizontal) and the other tilted 10° from the vertical (or horizontal) it made no difference to the size of the displacement which was made inspection- and which test-figure.[3] However, they used only one subject apart from themselves and, in view of the evidence cited above, suggesting that there is an independent Gibson-effect, the experiment seemed to be worth repeating on a larger scale.

Method: (1) Observers. The *Os* were 10 university students, most of them naïve in studies in this area.

(2) Apparatus and procedure. A split-beam tachistoscope was used, one channel carrying the inspection-figure and the other test- (and in Part II the comparison-) figure. Each stimulus consisted of an 11-in., 60-w. strip of light pivoted about its center and reduced to give a line of light 0.08 in. wide. Both test- and inspection-figures had small opaque fixation-points at their centers of rotation and were located in such a way that, when both figures were illuminated, their fixation-points coincided. *O* had his head clamped, and viewed the figures in a dark field from a distance of 40 in.

In each of the two parts of the experiment, a forced-choice, constant-stimulus method was used to establish a point of subjective equality or verticality, first under control conditions and later with the inspection-figure present. The control run consisted of 25 1.0-sec. presentations of the test-stimulus, five in each of five positions, in random order. The exposures were separated by 5.0 sec. with no figure visible during which the *O* responded that the line had appeared clockwise or anticlockwise relative to the vertical (in Part I) or to the comparison-figure (in Part II). After a few minutes' rest, *O* was required to fixate the inspection-line for 1.0 min., and this was followed immediately by the 25 presentations comprising the experimental condition. These were similar to the control run except that *O* now had to fixate on the inspection-line during the 5.0-sec. intervals separating the exposures of the test-stimuli.

In each part, the experimental condition always followed the control condition but the order in which the two parts were carried out was alternated among the ten *Os*. The two parts were separated by an interval of several days.

Results and discussion. In Part I, the inspection-line was tilted 10° clockwise and the five positions of the test-line ranged in half-degree steps from 1° clockwise to 1° counterclockwise. The usual after-effect of tilt, which is predictable from both theories, should therefore reduce the number of 'clockwise' responses in the experimental condition relative to the control condition, thereby indicating that lines close to the vertical appear turned away from the position of the inspection-figure.

In Part II of the experiment, the inspection-line was vertical and the five positions of the test-line ranged from 9° to 11° counterclockwise. In view of the greater difficulty of judging the precise orientation of tilted

[3] Wolfgang Köhler and Hans Wallach, Figural after-effects: An investigation of visual processes, *Proc. Amer. philos. Soc.*, 88, 1944, 4.

lines, the test-figure was always accompanied by a comparison-line, 4.0 in. to its left and tilted 10° counterclockwise. *O* had to say, after every exposure, whether the test-line appeared tilted clockwise or counterclockwise relative to the 10° comparison-line. In this situation, no displacement could be predicted from Gibson's theory, but a figural after-effect would produce an apparent displacement of the tilted test-line away from the position of the vertical inspection-line, *i.e.* fewer 'clockwise' responses in the experimental than in the control condition.

The total number of 'clockwise' responses made by every *O* in each condition was computed, and the means of these totals are set out in Table I.

It can readily be seen that in Part I the inspection-figure tilted clockwise produced the expected decrease in the number of clockwise responses

TABLE I

MEAN NUMBER OF 'CLOCKWISE' RESPONSES

Experiments	Part I	Part II
Control	14.5	9.4
Experimental	3.5	3.5
Difference	11.0	5.9

to the near-vertical test-figures. On the basis of a t-test for correlated means, this difference of 11 was found to be highly significant ($p < 0.001$).

Similarly Part II demonstrates the normal figural after-effect in which the tilted test-lines appear even more tilted after exposure of the vertical inspection-line. The mean difference of 5.9 is significant at the 1% level.

For the hypothesis of a pure tilt after-effect independent of the figural after-effect the displacement must be larger in Part I than in Part II. Again using the t-test for correlated means, we found the difference of 5.1 between the effects in the two parts to be significant at the 2% level.

Thus we conclude that, contrary to Köhler's claim, there is a mechanism of the sort proposed by Gibson, which operates when an inspection-figure is tilted and does not operate when it is vertical.

A possible criticism is that the use of a comparison-figure in Part II invalidates the direct comparison of the two parts. The similarity of the variances in the two cases suggests, however, that the tasks are in fact quite comparable. It could also be argued that the comparison-figure as well as the test-figure might be affected by the processes of the upright inspection-figure. Satiation, however, produced by the inspection-figure, while it would cause the test-line to appear more tilted than it really was, could only have the opposite effect on the comparison-figure. (Only in the special case where the test- and inspection-lines actually cross does the apparent angular separation increase; in other cases, that part of the test-figure which is

closer to the position of the inspection-figure suffers greater apparent repulsion, thereby reducing the apparent angular separation of the lines). This would make the measured effect in Part II larger relative to that in Part I, contrary to the hypothesis.

Use of a comparison-figure in Part I, on the other hand, is precluded by the probability that the Gibson-effect transfers to a considerable extent and would therefore tend to be partialled out through the inspection-figures having a similar effect on both test- and comparison-figures.[4] Indeed, Köhler and Wallach's use of a comparison-figure in the Part I, is the most likely explanation for their failure to obtain a difference between the two situations.

The response-count in Table I suggests that the two effects may be of similar magnitude since the numbers of responses representing the displacements in the two parts are in the ratio of almost 2:1. Unfortunately it is not possible to derive from the data a more precise estimate of their relative magnitudes based on the *P S Es*; the effects in both parts were so large relative to the range of the test-stimuli that the extrapolation required to estimate median points in the experimental conditions could only be the wildest guess.

In any case, evaluation of the effects by subtraction of this sort requires the assumption that they are additive, an assumption made extremely difficult to test by the apparent impossibility of devising a condition in which the Gibson-effect might be expected to occur alone.

Summary

Both Köhler's satiational theory and Gibson's adaptational theory can account for the well-known after-effect of a tilted line on the apparent vertical. The question of whether there are in fact two mechanisms was investigated by setting up a condition in which only the figural after-effect could be expected to occur; viz. a vertical inspection-figure and a tilted test-figure. Although an effect was obtained in this condition, it was significantly smaller than that found for the tilt after-effect. This was interpreted as evidence that both mechanisms operate.

[4] R. B. Morant and H. H. Mikaelian, Inter-field tilt after-effects, *Percept. mot. Skills*, 10, 1960, 95-98.

MULTIPLE PROBABILITY-LEARNING WITH SHIFTING WEIGHTS OF CUES

By Cameron R. Peterson, Kenneth R. Hammond, and David A. Summers, University of Colorado

Man's adjustment to an uncertain environment has been thoroughly studied in the classical experiment on probability-learning, in which the stimulus-situation is characterized by some degree of randomness but is stationary through trials. Uncertainty is associated with the event that occurs on any specific trial, but the probability of which event will occur remains constant over trials.[1]

Many uncertain situations in the real world are not only partially random, but are also nonstationary. Characteristics of such situations are not stable but change over time. A deviation at any time may be either a random fluctuation or the reflection of a genuine change in the situation. Adjustment to a nonstationary, uncertain environment requires that *S* detect whether a deviant event is a random fluctuation or the signal of a real change. Responses should ignore random fluctuations, but should react to genuine changes.

In the classical experiment on probability-learning, the probability of response stabilizes at, or somewhat above, the stationary probability of the corresponding event. When the probability of an event shifts as a function of trials, a corresponding shift occurs in the probability of the response.[2]

The present experiment is in the context of multiple probability-learning. Whereas previous experiments have compared the degree to which cue-weights (dependence upon cues) for responses correspond with

---

* Received for publication December 15, 1964. The research reported here was undertaken in the Behavior Research Laboratory, Institute of Behavioral Science, University of Colorado. The research was supported by a Research Grant (GM-9028) from the National Institute of Health.

[1] D. A. Grant, H. W. Hake, and J. P. Hornseth, Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *J. exp. Psychol.*, 42, 1951, 1-5.

[2] W. K. Estes, The statistical approach to learning theory, in Sigmond Koch (ed.), *Psychology: A Study of a Science*, 2, 1959, 412-413; see also M. P. Friedman, C. J. Burke, M. Cole, L. Keller, R. B. Millward, and W. K. Estes, Two-choice behavior under extended training with shifting probabilities of reinforcement, in R. C. Atkinson (ed.), *Studies in Mathematical Psychology*, 1964, 250-316.

stationary cue-weights of stimuli,[3] the purpose of the present experiment was to study the way in which weights of the cues for responses shift as a result of changes in objective weights of the cues.

Method: (1) Experimental design. Following the usual design of studies of multiple probility-learning, *S* on each trial observed the values associated with the cues, based his prediction of the criterion upon these values, and then observed the correct value of the criterion. The experimental problem required that the objective weights of the cues (*b*-coefficients) change as a function of trials. Accordingly, the three cues were assigned weights of ⅔, ⅓, and 0 for the first 100 trials. For the second 100 trials, of the 200 trial-task, weights were reversed to 0, ⅓, and ⅔.

(2) Stimulus array. The 200-trial stimulus-array was created by selecting at random 200 integers between 1 and 10 for each cue-variable. This resulted in rectangular distributions with approximately zero correlations between cues. Criterion-values were created by the following procedure: (a) each tentative criterion-value was set equal to the weighted sum of cue-values when each cue-value was multiplied by the weight (⅔, ⅓, 0) assigned to that cue; (b) random error, which reduced the stimulus-multiple R from 1.0 to about 0.90 was introduced by adding to each criterion-value a number randomly selected from a distribution having a mean of zero and standard deviation of one; and finally, (c) criterion-values were rounded off to the nearest integer between 1 and 10. The distribution of the resulting criterion-variable was approximately normal, with the middle values occurring about three times as frequently as the extremes.

(3) Subjects. The *Ss* were 29 students in nursing. They served in a single group and were not paid.

(4) Apparatus. A 3 × 4 ft. upright display-board permitted the values of the cues and criterions to be displayed to the *Ss*. The cues were represented by three scaled vertical columns labeled *A, B,* and *C*. The criterion was represented by a scaled row just below the three columns. Each column and the row had 10 equally-spaced pegs numbered from 1 to 10. *E* displayed the correct values of cue and criterion by hanging a ring over the appropriately numbered peg in each column and in the row.

(5) Procedure. On each of the 200 trials, *S*'s task was to predict as nearly as possible the correct value of the criterion. This prediction was to be based upon the cues. Each trial included three steps: (a) *E* displayed the three cues by hanging a ring on the appropriate number in each column of the display-board; (b) each *S* recorded his prediction in a booklet; and (c) *E* displayed the correct value of the criterion by hanging a ring on the appropriate number in the row of the display-board.

Non-technical instructions emphasized three aspects of the task: (a) it was partially random, hence it was impossible to predict correctly on each trial; (b) the criterion was a linear (weighted average) function of the cues; and (c) the

[3] S. A. Summers, The learning of responses to multiple weighted cues. *J. exp. Psychol.*, 64, 1962, 29-34; see also C. N. Uhl, Learning of interval concepts: I. Effects of differences in stimulus-weights, *ibid.*, 66, 1963, 264-273; and C. R. Peterson, K. R. Hammond, and D. A. Summers, Optimal responding in multiple-cue probability learning, *ibid.* (in press).

weights of the cues were not necessarily stationary; any change over trials in the relation between cues and criteria could be due either to a random fluctuation or to an actual change in cue-weights.

Results. The b-coefficients between the three cues and *S*'s responses measured the weighting of the cues by the responses of the *S*. This measure, calculated for individual *S*'s and then averaged across the *S*s, is shown as a function of blocks of trials in Fig. 1. The five pre-reversal blocks are designated by subscript *a* and the five post-reversal blocks by subscript *b*.

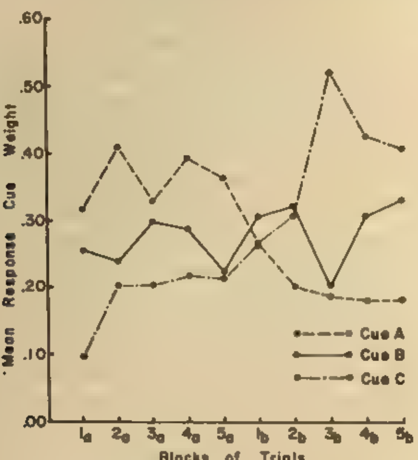

FIG. 1. MEAN RESPONSE *b*-COEFFICIENTS AS A FUNCTION OF BLOCKS OF TRIALS
The subscript *a* refers to trials before, and *b* to trials after the reversal in the *b*-coefficients of the stimuli.

Response-cue weights rank-ordered appropriately during block 1 *a*, and remained so during the first five blocks. Following the shift in the objective weights of the cues after Block 5_a, the b-coefficients for the responses also shifted, but did not achieve appropriate rank ordering until block. 3_b.

This description of Fig. 1 is supported by an analysis of variance. The cue-weights of responses corresponding to the different levels of cue-weights of stimuli were significantly ($F = 25.72$, $df = 2/56$, $P < 0.001$) separated both before and after the stimulus-shift. A triple interaction between level of cue-weight, blocks numbered 1-5, and blocks letered a-b ($F = 8.75$, $df = 8/224$, $P < 0.001$) indicates that the appropriate rank ordering of response cue-weightings was achieved at a different rate, *i.e.* more quickly, during the pre-reversal trials than during the post-reversal trials.

While not germane to this experiment, it should be noted that there were large individual differences in the degree to which the *b*-coefficients of responses for different *Ss* agreed with the objective weights. As is to be expected, those *Ss* who weighted the cues most appropriately were those who made the most accurate predictions of the criterion-values. For cues with theoretical weights of ⅔, ⅓, and 0, the corresponding mean subjective weights were 0.56, 0.36, and 0.11 for the *Ss* with the highest accuracy in each block, and −0.03, −0.02, and 0.18 for *Ss* with lowest accuracy. The *Ss* with intermediate accuracy yielded intermediate cue-weights nearer the means of Fig. 1.

Discussion. Since the rank order of the *b*-coefficients of the responses appropriately tracked the shift in the objective weights of cues, it is clear that *Ss* were able to respond to the nonstationary aspect of the task even in this rather complex situation. Nonetheless, two aspects of the results suggest that performance in a nonstationary task is more difficult than performance in a stationary one. First, although the means of the response *b*-coefficients showed an appropriate rank-order, they did not reach the objective values nor did they appear to be approaching them. Perhaps this lack of separation is partly due to instructions that specified that weights of the cue could change as a function of trials. There may have been a tendency for the *Ss* to interpret random fluctuations as actual changes in weights of the cues, and to change weightings of cues by the responses accordingly, even during periods when the objective weights of the cues did not change. Another possibility is that anticipation of a change resulted in a 'hedging of bets.' Anticipation of a change could have caused hesitancy in placing the appropriately strong weight on the cue with a weight of ⅔ and completely ignoring the cue with a weight of zero.

A second indication of the difficulty of a nonstationary task is that the tracking of the change in the weights of the cues was slower than the original learning. At the beginning of the experiment, there was nothing to unlearn. Only a response to the relative differences in cue-weights was necessary. After the stimulus-shift, it was necessary to detect that a change had occurred and to ignore the pre-shift weights of the cues while responding to the post-shift weights. These indications of difficulty point to the importance of studying nonstationary as well as stationary tasks in probabilistic situations.

ASSIMILATION AND ATTENTION IN VISUAL FORM-PERCEPTION

By Myra Adamthwaite and Juliet Popper Shaffer
University of Kansas

In a recent study, Gardner and Lohrenz found that the extent of assimilative interaction among visually-perceived forms presented sequentially was inversely related to the degree of attention paid to the forms during their presentation.[1] The authors compared the performance of control *S*s who viewed a series of paired figures under standard conditions with the performance of experimental *S*s who were required to count backwards by twos while viewing the stimuli. Reproductions of the figures were ranked for amount of assimilation, which was measured by deviations from the actual stimuli which could be attributed to the presence of the other stimuli.

It seemed to the present authors that the higher ranks on assimilation for the experimental group might have resulted from greater inaccuracy of all kinds due to decreased attention, or from other effects of decreased attention *per se,* rather than from an enhancement of assimilative interaction with other figures. In the present study, the *S*s reproduced only a single stimulus (referred to hereafter as the reproduction-stimulus or Figure *A*). One-third of the *S*s observed only this single stimulus (Group *A-A*), one-third observed this stimulus followed by another stimulus (Group *AB-A*), and one-third observed it preceded by another stimulus (Group *BA-A*). Differences in the assimilative scores of Groups *AB-A* and *BA-A* under different conditions of observation could then be compared with the corresponding differences in Group *A-A,* where no real assimilation could occur. The comparison of Group *AB-A* with Group *BA-A* also provided evidence as to the effect of order of presentation on assimilation.

In addition to the standard condition and the condition of distraction (counting backwards by twos) studied by Gardner and Lohrenz, the present study included a condition of speed (reduced stimulus-exposure time) to test the supposition of Gardner and Lohrenz that an externally-

* Received for publication February 20, 1965.

[1] R. W. Gardner and L. J. Lohrenz, Attention and assimilation, this Journal, 74, 1961, 607-611.

effected reduction of attention would have the same effect on assimilation as distraction, which presumably represented an internally-effected reduction of attention.

The experiment, then, consisted of obtaining reproductions of a single geometric figure under all combinations of three different 'stimuli presented' (*A-A*, *AB-A*, and *BA-A*) and three different 'Modes of presentation,' (standard, distraction, and speed). The hypotheses to be tested were: (1) Observation of two stimuli would result in higher assimilative scores than observation of only the reproduction-stimulus (*i.e.* true assimilation would occur); (2) the distraction- and speed-modes would result in higher assimilative scores than the standard mode of presentation (*i.e.* assimilative scores would be increased under conditions of reduced attention); and (3) the differences between assimilative scores under attention-reducing (distraction and speed) and standard modes of presentation would be greater in the *AB-A* and the *BA-A* groups than in the *A-A* group (*i.e.* the increase in assimilative scores under conditions of reduced attention would be due at least in part to an increase in true assimilation).

Subjects. The *Ss* were 108 undergraduates, 45 men and 63 women. They were distributed into the nine experimental groups according to a prearranged quasi-random schedule, with a total of 12 *Ss* per group. Each *S* was tested individually.

Materials. The pair of stimuli was Pair 4 in the study by Gardner and Lohrenz, and their Figure 4*A* was the reproduction-figure.[2] The *Ss* sat approximately 7 ft. from a wall upon which the figures were displayed by means of a slide-projector. The projections of both figures were 18½ in. in the longest dimension. The fluorescent room-light remained on during the experiment.

Procedure. The instructions for each group were similar to those used by Gardner and Lohrenz, with modifications necessitated by the presentation of only one pair of figures or one figure in this study. As in the previous study, exposure-time for all figures in the standard and distraction modes of observtaion was 0.5 sec., and there was a 2-sec. interval between exposures for the *Ss* presented with the pair of stimuli. An exposure-time of 0.075 sec. with a 2-sec. interval between exposures of the pair of stimuli was used for the speed-mode.

Since only a single stimulus was to be reproduced in this study, it seemed desirable to equate the time between exposure of Fig. *A* and its reproduction for all *Ss*. To that end, an interval of silence (ranging from 0–6 sec. according to the experimental condition) followed presentation of the stimuli, with the time between exposure and reproduction being 6 sec. under all conditions.

Measure of assimilation. In the study by Gardner and Lohrenz, three raters independently ranked all reproductions of each pair for degree of assimilation. In the present study, three raters independently scored each reproduction for degree of assimilation on a 7-point scale, high scores indicating high assimilation. The specific

[2] Gardner and Lohrenz, *op. cit.*, 608.

criteria which the raters were instructed to use were the same as those employed by Gardner and Lohrenz. Raters did not know what conditions had been used in the study nor that some of the *Ss* had seen only one figure.

Results. Pearson correlation-coefficients were computed between the sets of scores from each pair of raters. For Raters 1 and 2, the correlation was 0.91; for Raters 1 and 3, 0.87; and for Raters 2 and 3, 0.95. Since the agreement among raters was so high, the median score for each reproduction was used in the analysis of the data.

The mean assimilation-scores for the different experimental groups are presented in Table I. An analysis of variance was carried out with 'Stimuli presented' and 'Mode of presentation' as factors. The effect of Stimuli was significant ($F = 4.47$, $df = 2/99$, $p < 0.05$), and the differences were in the predicted direction. The difference between the means

TABLE I

MEAN ASSIMILATION-SCORES FOR THE EXPERIMENTAL GROUPS

Stimuli presented	Mode of presentation			
	standard	speed	distraction	combined
A-A	2.83	3.57	4.25	3.55
AB-A	3.82	4.50	5.07	4.46
BA-A	3.50	5.32	5.82	4.88
Combined	3.38	4.46	5.05	—

of the *A-A* group and the *AB-A* group was of borderline significance ($t = 1.92$, $df = 66$, $p = 0.06$), and the difference between the means of the *A-A* group and the *BA-A* group was significant ($t = 3.00$, $df = 66$, $p < 0.01$).[3] The difference between the means of the *AB-A* group and the *BA-A* group was not significant ($t = 0.93$).

The effect of mode of presentation also was significant ($F = 6.89$, $df = 2/99$, $p < 0.01$), with all differences in the predicted direction. The difference between the means of the standard and distraction-groups was significant ($t = 3.77$, $df = 66$, $p < 0.01$), as was the difference between the means of the standard and speed-groups ($t = 2.26$, $df = 66$, $p < 0.05$). The difference between the means of the distraction- and speed-groups was not significant ($t = 1.31$).

The interaction between stimuli presented and mode of presentation fell far short of statistical significance ($F = 0.41$). This result does not support the third hypothesis; namely, that differences between standard and attention-reducing modes of presentation would be greater in the

[3] Every t-test reported was carried out using variance estimates based upon the scores of only the groups involved in the test.

AB-A and the *BA-A* groups than in the *A-A* group. The observed differences actually are smaller in the *AB-A* group than in the *A-A* group. Nevertheless, since the over-all *F*-test for interaction is not highly powerful as a test for the particular type of interaction hypothesized, two further analyses were performed: the differences between standard and distractive modes of presentation were compared for the *BA-A* and the *A-A* groups, and the differences between standard and speed-modes were compared for the same two groups. Neither comparison approached statistical significance ($t = 0.87$ and 0.94, respectively, $df = 44$).

Discussion. Confirmation of the first hypothesis indicates that assimilation did occur under the given experimental conditions: assimilative scores for drawings of the reproduction-stimulus were significantly higher when the additional stimulus was presented than when it was not. The second hypothesis; namely, that assimilative scores would be higher under distraction and speed-modes of presentation than under standard conditions also was fully confirmed, and the results of this study therefore are consistent with those of Gardner and Lohrenz. As already noted, however, this increase in assimilative *scores* under conditions of reduced attention is not necessarily evidence that assimilation itself was greater, since other variables could account for the higher scores. In the test of the third hypothesis, the results indicated that assimilative scores increased just as much when attention-reducing models of observation were introduced if only a single figure had been observed (and thus no true assimilation could occur) as if the pair of figures had been observed. The over-all outcome of the present study, then, provides no support for the hypothesis that assimilation is increased under conditions of reduced attention.

POSITION-CUES IN SERIAL LEARNING

By Donald L. McManis
Bureau of Child Research, University of Kansas

Ebenholtz pointed out that a serial list may be mastered by learning the location of the items relative to the beginning and end of the list, by forming sequential associations between successive items, or on the basis of both processes.[1] He conducted an experiment in which one group learned a 10-item nonsense-syllable list in the usual manner, with the words always appearing in the same serial position, whereas a second group began each trial at a different point in the sequence. The latter procedure produced significantly greater difficulty, even though the sequence and number of associations were identical in both conditions. This experiment yielded evidence for the role of positional cues in serial learning as opposed to a solely sequential association explanation of the serial learning process.

Winnick and Dornbush investigated the role of positional cues in serial learning, using three experimental conditions: a constant starting position: a starting position shifted by one syllable on successive trials; and a starting position randomly varied.[2] Their findings supported Ebenholtz's conclusion that positional cues play an important role in the serial learning.

In these studies, nonsense-syllables of high associative value, but with no demonstrated relationship of inter-item stimulus-response, were used. It may be the case that sequential associations are more difficult to form between such items than between actual words. It would be a stronger test of the positional learning if it could be demonstrated that removal of the positional cues made serial learning more difficult even with words having a known stimulus-response relationship.

With this in mind, in the present study, an attempt was made to enhance the probability of sequential associations being formed by the selection of the serial list. In addition to using *CVC* trigrams of high

* Received for publication March 18, 1965. This investigation was conducted at the Parsons State Hospital and Training Center, Parsons, Kansas, and supported with funds provided under NIH Grant NICHHD 00870-01.

[1] S. M. Ebenholtz, Serial learning: Position learning and sequential associations, *J. exp. Psychol.*, 66, 1963, 353-362.

[2] W. A. Winnick and R. L. Dornbush, Role of positional cues in serial learning, *ibid.*, 66, 1963, 419-421.

associational value, lists were formed of words so arranged that each was preceded by another one which had been found to elicit it in a test of free-association. *S*s learned these lists in the regular manner, with the words always having the same serial position, and with the starting point and serial positions being randomly varied from trial to trial.

Method: (1) Subjects. The *S*s were 24 students at Parsons Junior College in Parsons, Kansas. They were paid to participate in the study.

(2) Materials. Four basic 10-item lists were used. Tow lists were composed of 87% association-value trigrams from Glaze's list: (1) HOV, CAK, BIC, GUZ, DEB, FOM, LUR, KAS, MID, and JEN; and (2) HUF, CER, BUX, GOL, DAW, FIJ, LEM, KIP, MOS, and JAV. The other two lists consisted of words from the Kent-Rosanoff word list, so arranged that each item was an empirically demonstrated stimulus-word for the one that followed it:[8] (3) Thirsty, Hungry, Stomach, Eating, Sleep, Bed, Soft, Hard, Working, and Man; and (4) Swift, River, Deep, Ocean, Blue, Red, White, Dark, Light, and Lamp. The lists were each so arranged with four different randomly determined starting points that the list began with a different item, but maintained the same item-sequences, from trial to trial. In each list the initial three items were repeated at the end of the list.

(3) Procedure. The lists were learned by the anticipation-method, at a 2-sec. rate with a 14-sec. inter-trial interval. Every *S* performed in all four procedures (Trigram-Constant, Trigram-Varied, Word-Constant, and Word-Varied) with the procedural order being controlled by a Latin square. The *S*s learned different lists in each procedure and each of the four lists was learned equally often in the varied and constant conditions. The *S*s were assigned to a specific procedure in the order they appeared at the laboratory.

In the constant, starting-point procedure, the *S*s were told that they would be shown some words (or letters) which would appear in the window of the machine (a Lafayette model 303-B memory-drum), and that they were supposed to "learn what item followed every other item. They were told to say the word (or spell the trigram) that came next each time an item appeared.

In varying the starting point, the *S*s were given the same instructions, but were also informed that the list would begin with a different item each trial. It was pointed out that the initial three items would allow them to get into the sequence, and that these items were repeated at the end of the list.

The *S*s were given 10 anticipation-trials in each procedure. The criterion of learning used was total errors made during each procedure. The responses made to the initial three items in each list were not counted in the total errors, since those items were designed to get *S* into the rhythm and were repeated at the end of the list for the measure of response.

---

[8] E. Z. Rothkopf and E. U. Coke, Intralist association data for the 99 words of the Kent-Rosanoff word list, *Psych. Rep.*, 8, 1961, 463-474.

Results. The relative effects of the constant and varied starting positions were evaluated by comparing the total errors in each procedure. The error-scores were transformed by a square root transformation due to heterogeneity in the variances ($F_{Max.} = 15.92$; $P < 0.01$). The means and *SD*s of the transformed error-scores are presented in Table I.

A Treatments × *Ss* analysis of variance revealed that the main effect of treatment was significant beyond the 0.01 level ($F = 62.87$, $df = 3/69$).[4] Since the treatment of primary concern was the constant vs. varied starting position, the main treatment was further analyzed into its three component parts (trigrams vs. words; constant vs. varied starting position; and the interaction between the type of material and starting position), each having one degree of freedom. After the variance associated with the type of material and the interaction between material and starting position had been removed, the variance associated with the constant vs. varied starting position was found to be significant beyond the 0.01 level ($F = 25.28$, $df = 1/69$). The variance associated with type of material (trigrams vs. words) was also significant ($F = 162.43$, $df = 1/69$), but the interaction between materials and starting position was not ($F = 1.03$ $df = 1/69$).

TABLE I

TOTAL ERRORS IN FOUR PROCEDURES OVER 10 TRIALS

Tirgram-Constant		Trigram-Varied		Word-Constant		Word-Varied	
mean	*SD*	mean	*SD*	mean	*SD*	mean	*SD*
5.91	1.47	6.93	1.63	2.43	1.07	3.96	1.12

Comparisons between constant and varied starting position within the same list types (trigram constant vs. trigram varied; word constant vs. word varied) were carried out using Lindquist's critical difference-procedure.[5] With both the trigrams and words, the varied starting position treatment resulted in significantly more errors than the constant starting position treatment (critical difference at the 0.005 level with 46° of freedom = 0.96).

Discussion. The findings of Ebenholtz and of Winnick and Dornbush were replicated in the present study by the trigram-data. The results of the word-list are viewed as strengthening the conclusion that positional

[4] E. F. Lindquist, *Design and Analysis of Experiments in Education and Psychology*, 353-362.

[5] Lindquist, *op. cit.*, 90-96.

cues play an important role in the serial learning process. Since the word-lists were composed of items having a known stimulus-response relationship between adjacent items, the learning of these lists without relying on positional cues should have been greatly facilitated. The finding that removal of positional cues resulted in a significant decrement in learning, even with such highly inter-related words, indicates that explanation of the serial learning solely on the basis of sequential associations is inadequate.

APPARATUS

A SYSTEM FOR RECORDING THE ORIENTING REFLEX: A PHOTOELECTRIC PLETHYSMOGRAPH AND TRANSISTOR AMPLIFIER

By PAUL F. GRIM and ROGER A. ANDERSON, The University of Chicago

The apparatus described here consists of a DC transistor amplifier, capable of driving most recording galvanometers, and a photoelectric plethysmograph for the finger and forehead which is attached to the amplifier.[1] The amplifier open-loop current gain is 65,000. The application of negative feedback proportional to galvanometric current insures excellent linearity at all frequencies well below galvanometric resonance. Use of batteries and transistors as circuit elements in a balanced bridge output stage results in an amplifier with low current consumption and drift. Parts for an amplifier and transducer cost less than $35.

The apparatus was constructed to study the vascular component of the orienting reflex, which is an organism's normal response to any novel stimulus or change in stimulation. The orienting reflex is thought to mediate all sensory preconditioning and to serve as a central integrator of the stimuli in classical conditioning. Russian investigators have suggested that a constriction of digital blood vessels accompanied by a dilatation of those in the forehead is always a sign of the orienting reflex.[2] Repeated presentation of a stimulus which produces an orienting reflex leads to habituation, while an organism's vascular response to a hot or cold stimulus, which is thermoregulatory, does not habituate and is known as the adaptive reflex. A stimulus too strong for normal functioning of the organism produces vascular constriction in the forehead as well as the finger and is known as a defensive reflex.

Fig. 1 presents the circuit diagram of the plethysmograph and amplifier. For the purpose of studying the orienting reflex it is necessary to build two such units which are similar electrically but differ physically in the construction of the transducer.

The transducer consists of a CL 505L Clairex photoconductive cell inserted in one arm of a Wheatstone bridge. The peak spectral response of this particular cell is 5500 Angstroms, in the yellow-green band, thus reducing sensitivity to blood coloration. Since a photo cell changes in resistance as the light impinging upon it

[1] Development of this equipment was done under USPHS Grant M-3639 to Sheldon H. White, who directed this research.

[2] Gregory Razran, The observable unconscious and the inferable conscious in current Soviet psychophysiology: Interoceptive conditioning, semantic conditioning, and the orienting reflex, *Psychol. Rev.*, 68, 1961, 81-146.

changes, blood volume changes are converted into changes in electrical resistance by directing light from a G-E Miniature Lamp No. 222 through the finger tip to the cell. In the case of the forehead, the bulb and photo cell are placed side-by-side instead of facing each other as with the finger unit. With both bulb and cell facing in the same direction and a small partition of about 1/8 in. between them, light must pass into the skin and diffuse across the partition to reach the cell. The lamps are connected to a 1.5-v. battery which is not shown in the schematic diagrams.

Fig. 2 illustrates the transducers. Dimensions are not critical and almost any opaque material can be used for construction. The lamp in the forehead unit should be tilted slightly toward the photo cell, which is perpendicular to the skin. The forehead unit as illustrated was constructed of wood drilled to accommodate the lamp, cell and wires. Two Formica plates screwed to the block serve as covers.

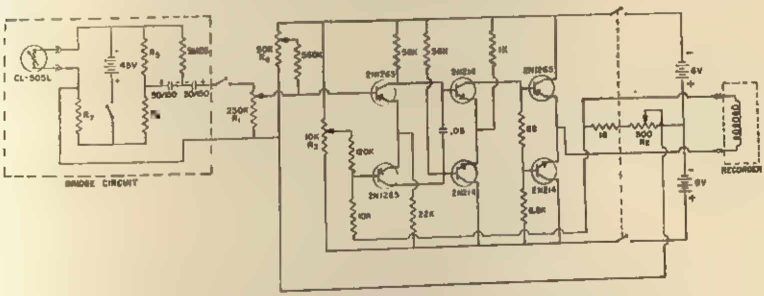

FIG. 1. SCHEMATIC DRAWING OF ONE TRANSDUCER BRIDGE-CIRCUIT AND AMPLIFIER.

The Two 56 K Resistors in the Top Center Were Modified to Two 5 K Resistors with Improved Amplification Symmetry.

The assembled unit is glued to a felt cap which cushions the forehead and protects the photo cell from changes in ambient room illumination. A small elastic belt holds the transducer against the forehead. The finger stock is simply an adjustable ring containing the photo cell and lamp. As shown in the sketch, both amplifiers and batteries can be housed in a small portable enclosure.

To determine the resistance values used in the bridge (R_5, R_6, R_7), find the average resistance of the photo cell while the transducer is in position on a subject and the small bulb turned on as in actual use. The resistance of the photo cell under this condition is the value for R_7. R_5 and R_6 should be equal and may be near the value of R_7. The authors found the following values satisfactory: Finger bridge, R_5=22 K, R_6=22 K, R_7=18 K. Forehead bridge, R_5=33 K, R_6=33 K, R_7=25 K.

Electrolytic capacitors were used for coupling because of their low cost. A non-polarized capacitor could be substituted in which case the 3-meg. resistor should be removed. By coupling the photo cell to the amplifier with a capacitor, blood-volume changes rather than absolute blood-volume units are recorded. When the resistance of the photo cell is not changing, the pen of the recorder will return to base line position no matter what the resistance of the cell. The large capacitors used for coupling pass very slow vascular changes. A slow change of 4-sec. duration will

produce an almost full scale deflection of the pen. As is to be expected, the sudden introduction or withdrawal of the finger from the transducer will produce a very large deflection requiring 15 sec. or more for return of the pen to the base line.

In actual use the pen is constantly deflected about 10% of its total travel by peripheral pulse wave-forms. Superimposed on the relatively rapid pulsations is a slower wave representative of relative vascular volume changes. Fig. 3 illustrates a typical plethysmogram.

Potentiometer R_1 is used to adjust the sensitivity of the Amplifier R_2 which can

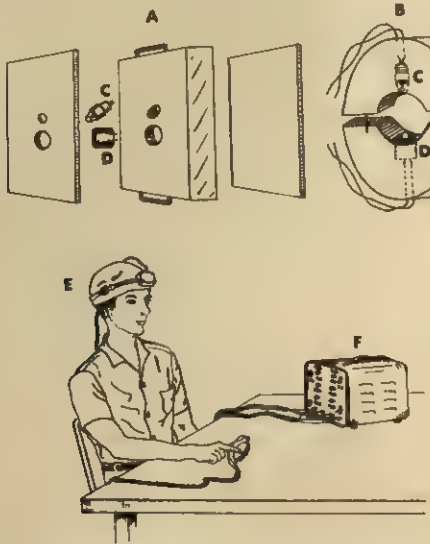

FIG. 2. THE TRANSDUCERS
(A) Enlarged view of the forehead-transducer; (B) Finger-transducer; (C) Lamp; (D) Photocell; (E) Subject with transducers in recording positions; (F) Enclosure containing two bridge-circuits, amplifiers, and batteries.

be a screw adjust potentiometer, will also reduce sensitivity while increasing negative feedback. Since negative feedback improves the linearity and input impedance of the amplifier at the cost of sensitivity, R_2 is a compromise adjustment. In actual practice, R_2 can be initially adjusted so as to provide just adequate sensitivity for most occasions and then locked out of use. Potentiometers R_3 and R_4 are used to zero set the recorder pen. Use of a balanced bridge output stage eliminates the need for a center-tapped pen. When the circuit elements are balanced, no current flows through the pen, while any imbalance, in either a positive or negative direction, causes a corresponding deflection of the pen.

The amplifier does not include temperature compensation, except for that inherent in the differential connection of the first two stages. With room tempera-

ture reasonably stable, however, the systems drift from zero-position during a 10-min. test-period was almost imperceptible. Considering that base line values are

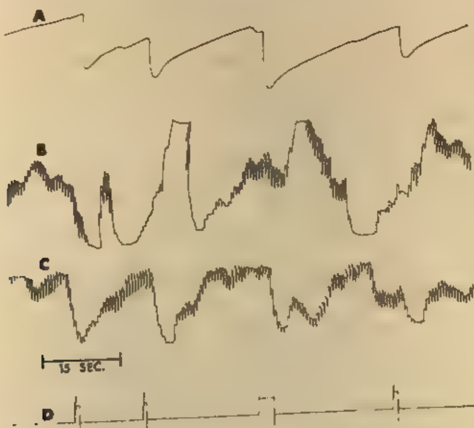

FIG. 3. RECORDING ILLUSTRATING A SERIES OF ORIENTING REFLEXES TAKEN WITH DESCRIBED INSTRUMENTATION

(A) GSR.; (B) Forehead Tracing; (C) Finger Tracing; (D) Stimulus Event Marker. The stimulus consisted of the click made by the event marker switch. Subject was instructed to sit in a relaxed position with his eyes closed. In both B and C an upward movement represents dilatation and a downward movement represents constriction. In B the limited travel of the recorder pen restricted two upward deflections.

TABLE I

ELECTRICAL TEST-RESULTS ON AMPLIFIER*

Current gain (open loop)	65,000 (96.2 db)
Voltage gain (open loop)	550 (54.8 db)
Linear output current (maximum)	10 ma. P-P
Linear output voltage (maximum)	6 v. P-P
Maximum current gain (closed loop)	40,000 (92.0 db)
Maximum transconductance (closed loop)	50,000 umho
Input impedance (open loop)	13,000 ohms
Input impedance at maximal gain (closed loop)	80,000 ohms
Input impedance at minimal gain (closed loop)	150,000 ohms
Linear frequency-response (5,000 Ω resistive load)	d-c to 100 ∼
Frequency-response (voltage gain 100 and resistive load)	−3 db. at 6.5 kc.
Frequency-response (voltage gain 10 and resistive load)	−3 db. at 50 kc.

* Transducer bridge circuit disconnected.

not being recorded, this does not seem objectionable. Actual test-results on the amplifier with the bridge circuit disconnected are given in Table I.

Adjustment procedure. R_1 is turned to minimal sensitivity. R_4 is then adjusted

to zero the recorder. R_1 is turned to about 50% sensitivity and the recorder pen is again adjusted to zero, this time using R_4. Thereafter, R_4 is not used and all further centering adjustments are made with R_3.

Use of the amplifier need not be limited to a plethysmograph or capacitor coupling. Applications include amplification of strain gage bridges, resistance bridge and other low-level transducer signals to drive recording galvanometers and other systems of acquiring data.

APPARATUS NOTE

A SIMPLE PROCEDURE FOR MAKING SMALL ELECTRODES

Many devices have been described to aid in making electrode assemblies for use in central stimulation.[1] A simple method of rapidly constructing small electrode units is to cast the component parts in a sipping-straw mold. A one-half inch segment of straw is filled with moist acrylic. Insulated electrodes are inserted in the center of the mold. A stiff wire is placed in the top of the mold to hold the unit in the electrode carrier of the stereotaxic instrument. When dry, the paper can be removed from the hardened dental cement. The unit is immediately ready for implantation and needs no further work such as edge trimming, filing to size, etc. This procedure helps to cut costs by conserving both acrylic powder and solvent as well as by eliminating miniature electrical plugs.

Central Washington State College

JOHN A. BULL, III
THOMAS B. COLLINS, JR.

[1] G. C. Sheatz, Electrode holders in chronic preparations, Part A. Multilead techniques for large and small animals, in D. P. Sheer (ed.) *Electrical Stimulation of the Brain,* 1961, 45; N. E. Miller, E. E. Coons, Melissa Lewis, and D. D. Jensen, Technique for use with the rat, *ibid.,* 51.

NOTES AND DISCUSSIONS

A REINTERPRETATION OF THE NATSOULAS AND DUBANOSKI STUDY

The purpose of the present paper is not to criticize the Natsoulas and Dubanoski study, but to reinterpret its results.[1]

Through an elaboration of a technique suggested by Krech and Crutchfield,[2] Natsoulas and Dubanoski attempted to infer the locus and orientation of a perceiver under various experimental conditions. The essence of their study is presented below in their summary statement.

> A letter is drawn on *S*'s head: on the forehead; the right or left side; or the back of his head. This letter may be perceived from an *external* locus *i.e.* from *E*'s vantage point, or as its *mirror image* from an *internal* locus. The conceptual scheme presented offers two intermediary causes of how the letter will be perceived: (a) the orientation of the perceiver; and (b) the magnitude of change in orientation necessary for perceiving the letter from an internal locus. The greater the shift necessary, the greater the likelihood of perception from an external locus.

The orientation of the perceiver is assumed to be forward when *S* is facing forward. When *S*'s head is turned right or left, the perceptual orientation may correspond to the head-direction, to the frontal direction, relative to the body, or to a direction between the latter two.[3]

The results support the conceptual scheme. They also suggest that the orientation of the perceiver is at an angle between head-orientation and the frontal direction.[4]

The authors of the article under consideration posit a perceiver who can assume various orientations both inside and outside the physical body. With the head forward, the locus of the perceiver is assumed to be internal (inside looking out) and the orientation forward. When *b* is traced on *S*'s forehead, it will be experienced as *d*. To experience the figure as *b* would require shifting the locus of the perceiver and would require greater "perceptual effort" than maintaining the current locus. If *b* is traced on the back of *S*'s head, according to the principle of perceptual effort, *S* should experience the *b* as *b* (*i.e.* from an external locus). To

---

[1] Thomas Natsoulas and R. A. Dubanoski, Inferring the locus and orientation of the perceiver from responses to stimulation of the skin, this JOURNAL, 77, 1964, 281-285. Throughout the present paper *b* and *d* are referred to as examples of letters traced on *S*'s head; in fact, Natsoulas and Dubanoski used *d*, *p*, *b*, and *q* in repeated sequence such that all the *Ss* received 12 tracings in each condition of the experiment.

[2] David Krech and R. S. Crutchfield, *Elements of Psychology*, 1958, 205.

[3] Natsoulas and Dubanoski, *op. cit.*, 285. Italics are mine.

[4] Natsoulas and Dubanoski, *op. cit.*, 285.

experience the *b* as *d* on the back of the head would require a shift in the orientation of the perceiver by a magnitude of 180° inside the head and, hence, would require greater perceptual effort than a change in locus.

On the basis of the principles posited, other predictions with respect to internal and external loci of the perceiver were made concerning tracings on the forehead, right and left temples, and the back of the head when the head is forward, turned right, and turned left. The predictions are neatly confirmed by their data. The essence of their results is presented in Table I.

A phenomenological reinterpretation. A phenomenological analysis of the task presented to their *S*s by Natsoulas and Dubanoski reveals that these authors have imposed certain assumptions on their data which, in fact, may have prejudiced their interpretations. One of these assumptions

TABLE I

PROPORTION OF *S*S SHOWING INTERNAL PERCEPTUAL LOCUS
(From Table I of Natsoulas and Dubanoski)

Orientation of head	Place where letter traced on head			
	forehead	right side	left side	back
facing forward	.70	.47	.49	.12
turned right	.69	.29	.56	.18
turned left	.62	.57	.36	14

is a mobile perceiver which is relatively independent of the physical body. Another concerns the notations of "inner" and "outer" which, as Straus points out, are not strictly spatial concepts.[5] A third assumption is the principle of perceptual effort. An additional assumption is that, experientially, a *d* is the mirror image of a *b*.

By setting these assumptions aside, it is possible to gain a clearer view of *S*'s experiences of the stimulus-letters used. For example, to say that *d* is the mirror image of *b* is to focus on the physical aspects of these stimulus-letters and to bypass the manner in which they are experienced. That is, *d* is not *experienced* as the mirror image of *b*. It is only on a derived, logical level that *d* and *b* are called mirror images. When a *b* is experienced, it is experienced immediately and directly as *b* and not by reasoning that it is the mirror image of *d*.

Asking now, what is the essence of a *b*-stimulus, it must be answered that, among other aspects, it is a Gestalt oriented to the right. Furthermore, a *d*-stimulus is essentially oriented to the left. A *b*-stimulus, then, does not occur in itself, but is a *b* for me—it is a figure (on a ground) occurring in my oriented or lived space. This leads directly to a consideration of a fundamental discovery by Merleau-Ponty, the body-subject.

[5] Erwin Straus, *The Primary World of Senses*, 1963, 241-247.

According to Merleau-Ponty's phenomenological analyses, the body itself is a meaning-giving existence—a subject. The oriented space in which our life runs its course is, to the conscious subject, *already* oriented. That is, I am already bound by certain dimensions which I simply have to accept. In this oriented space things are for me far-away or close-by, high or low, right or left. High or low and right or left refer to space only as it is for a subject. The subject for which such space exists is not the conscious and free subject but the body-subject.[*]

With respect to our lived body, then, things are right or left, high or low; these dimensions are constituted by the body-subject on a not yet conscious, not yet free level of existence. On the conscious level, these already constituted dimensions of oriented space are merely given. In

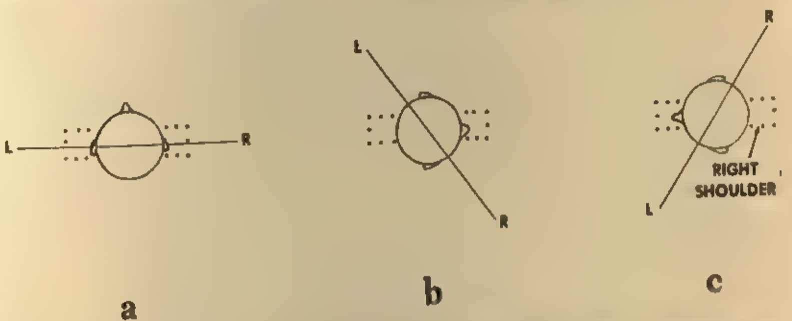

FIG. 1. RIGHT-LEFT BODY ORIENTATIONS
(a) head straight forward; (b) head turned to the right; (c) head turned to the left.

terms of oriented space, a *b* is oriented to the right and a *d* to the left. Thus, by calling on Merleau-Ponty's findings alluded to above, the results of the Natsoulas and Dubanoski study can be reinterpreted solely in terms of a bodily constituted right-left orientation. Specifically, since, from the point of view of the body-subject, a *b* is oriented to the right and a *d* to the left, when a *b* or *d* tracing is opposed to the right-left orientation of the body, they will be clearly discriminated; when *b* and *d* are not opposed to this body orientation, they will be confused.

When the head is facing forward, the right-left orientation of the body can be expressed by a line extending from ear to ear as in Fig. 1 a.

[*] For an extended discussion of the relationship between the body-subject and oriented space see: Maurice Merleau-Ponty, *The Phenomenology of Perception,* 1962, 98-147, 243-298; R. C. Kwant, *The Phenomenological Philosophy of Merleau-Ponty,* 1963, 21-23; Joseph Kockelmans, Merleau-Ponty's View on space-perception and space, *Review of Existential Psychology and Psychiatry,* 4, 1964, 69-105.

Consequently, when *E* traces a *b* on *S*'s forehead, *S* will experience the traced Gestalt as a *d* in terms of his body orientation (the Gestalt is oriented to his bodily left). On the other hand, when *E* traces a *b* on the back of the head, *S* will experience the traced Gestalt as a *b* (in terms of his body orientation, the traced Gestalt is oriented to the right). If *E* traces a *b* and a *d* on either the right or left temple, *S* will confuse them because the traced Gestalt will have a front or a back orientation with respect to the body, and this orientation does not correspond to the 'normal' left or right orientation of a *d* or a *b*.

On the basis of their results, Natsoulas and Dubanoski infer that when *S* turns his head to the right (left), his forward orientation shifts to a direction which falls somewhere between the 'normal' forward and far-right (far-left) orientation. The right-left orientation is, of course, correspondingly displaced. A more direct approach seems to confirm the above inference. If an *S* is asked to close his eyes, turn his head to the right, and then point to the left, it will be noted that *S* points somewhere *between* the normal forward and far-left axis. Corresponding results are obtained when *S* is asked to turn his head to the left. Thus, when the head is turned to the right, the left-right axis can be represented by a line extending from behind the left ear to in front of the right ear as in Fig. 1 b. When the head is turned to the left, the left-right axis extends from in front of the left ear to behind the right ear as in Fig. 1 c.

It will be noted from Table I that Natsoulas and Dubanoski find, consistent with their predictions, that when the head is turned to the right (left), there is a greater frequency of an internal locus of the perceiver when the *b* is traced on the left (right) temple in comparison with when it is traced on the right (left) one. These results present no problem to the theory of a bodily-constituted oriented space presented above. From this point of view, when the head is turned to the right, the right-left orientation of the body now extends in a manner illustrated in Fig. 1 b. Now, when *E* traces a *b* on the left temple, it is experienced as a *d* (the traced Gestalt is oriented to the left with respect to the 'displaced' right-left axis of body orientation). For the same reason, *b* traced on the forehead is also experienced as a *d* (oriented to the left). On the other hand, when *E* traces a *b* on the right temple it will be experienced as a *b* (oriented to the right with reference to the shifted right-left orientation of the body). For similar reasons, a *b* traced on the back of the head will also be experienced as a *b*.

Following the same line of reasoning, when the head is turned to the left, the *b*s traced on the forehead and right temple will be experienced as *d*s, and the *b*s traced

on the left temple and back of the head will be experienced as *b*s. These expectations are corroborated by the results presented in Table I.

From the point of view based on Merleau-Ponty's thought, it becomes clear that *S* consistently uses the *same* orientation to judge the experienced stimuli, and this orientation is a bodily-constituted oriented space. The notions of internal and external loci arise when *E* changes *his* position such that the *relationship* between his own body orientation and *S*'s body orientation changes. Thus, when *E* judges *S*'s reported experience in terms of his own (*i.e.* *E*'s) oriented space (which *E* is fully aware *he* does not change; after all, the *b* looks like a *b* to him no matter where he stands with respect to *S*), it *appears* that *S* must be changing his perceptual locus. This appearance is a consequence of *E*'s imposition of his own point of view or body orientation on *S*. This point will now be clarified by considering a recent study by Dubanoski.[7]

In the second experiment of the Dubanoski study, *E* stimulates only the top of the head while taking up various positions with respect to *S* (in front, to the right, to the left, and behind). The results indicate that when *E* stands in front of or in back of *S*, the proportion of internal perceptions corresponds roughly to the Natsoulas and Dubanoski results. Furthermore, when *E* stands to the right or left of *S* when tracing on the top of the head, there are *more* internal perceptions than occur when *E* stands in front of *S*. As Dubanoski finally admits, "the conceptual scheme used in Experiment I [and in the Natsoulas and Dubanoski study] to explain this effect, therefore, cannot be used here."[8] Dubanoski further states that "more empirical evidence is needed before a theoretical model can be generated which would subsume the results of the present experiments [and those reported by Natsoulas and Dubanoski]."[9] What is needed is *not new evidence* but rather a *new way of looking* at the available data. This is precisely the purpose of the present paper.

The task is now to account for the Dubanoski results in terms of the point of view provided by Merleau-Ponty. The first result of interest is that *S* consistently (about 80% of the time) agrees with *E*'s point of view when *E* stands behind *S* (and their body orientations correspond), and *S* consistently disagrees (about 80% of the time) with *E*'s point of view when *E* stands in front of *S* (and their body orientations do not correspond with each other). That is, *S* consistently experiences the stimulus-Gestalt with reference to his *own* body orientation.[10]

When *E* traces on the top of *S*'s head while standing to the right or

[7] R. A. Dubanoski, Inferring the phenomenal locus of the perceiver from responses to tactile stimulation, Unpublished Master's thesis, University of Michigan, 1964.

[8] Dubanoski, *op. cit.*, 23.

[9] Dubanoski, *op. cit.*, 25.

[10] Dubanoski did not use letters as stimulus-Gestalts. For simplicity in illustrating our points, reference is made to the letters *b* and *d*.

left of *S*, it is noted that even fewer external perceptions are reported than occur when *E* stands in front of *S*. This result is not intelligible in terms of the conceptual scheme proposed by Natsoulas and Dubanoski. The results of Dubanoski's study are accounted for by the point of view presented in the present paper. Fig. 2 will help the reader grasp our understanding of these results.

From Fig. 2 a, it will be noted that the body orientations of *S* and *E* are in correspondence. Thus, when *E* imposes his own orientation on *S*, the correspondence allows both *E* and *S* to experience the stimulus-object as a *b*. Hence, *E* judges that *S* is viewing the stimulus from an external locus; namely, from *E*'s *own* point of view.

When *E* stands in front of *S* (Fig. 2 b), *E*'s right-left orientation does not correspond to that of *S*. Hence, when *E* experiences that he draws a *b* (a stimulus-Gestalt

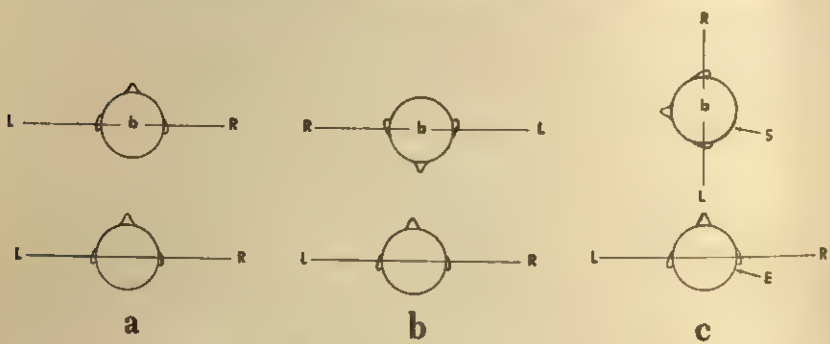

Fig. 2. Relationships Between the Right-Left Body Orientation of *E* and *S* (a) when *E* stands behind *S*; (b) when *E* stands in front of *S*; and (c) then *E* stands to the left of *S*.

oriented to his right) on *S*'s head, *S* experiences a *d* or a *q* (a stimulus-Gestalt oriented to his left). *E*'s judgment, then, is that the locus of the perceiver is other than his own (*i.e.* internal). By comparing Fig. 2 a and Fig. 2 b, it can be seen that for *E* the stimulus-Gestalt remains the same, but for *S* it has changed from a right to a left oriented Gestalt. The perceptual locus of neither *E* nor *S* has changed! The only change is in the *relationship* between the orientation of *E*'s body and that of *S*'s.

When *E* traces a *b* on top of *S*'s head while standing to his left (or right), it is virtually impossible for *S* to spontaneously experience the stimulus-Gestalt as a *b* (see Fig. 2 c). If a *b* is essentially a stimulus-Gestalt with a low-right orientation, *S* will not be able to experience the traced Gestalt as a *b*. With respect to *S*'s body-orientation, the stimulus will be experienced as a *q* turned sidewise. Since *E* will only give 'credit' to *b* as an external perception, it is not surprising that such a perception virtually never occurs when *E* stands at either side of *S*.

At least with respect to the conditions in which *E* traces on top of *S*'s head from the front or back (or when *E* traces on *S*'s forehead or back of head), 100% con-

sistency of internal or external perception is the theoretical expectation. In view of the repeated finding that this consistency distributes itself around values of approximately 70 to 80%, 20 to 30% may be taken as the 'normal' proportion of the time that stimuli are confused in this type of experimental situation. When *E* stands to *S*'s right or left and traces on top of the head, apparently there is nearly 100% confusion and not 100% internal perception.

Conclusions. By considering the Natsoulas and Dubanoski experiment from the point of view of phenomenology, it was possible to by-pass several assumptions made by these authors and to penetrate to the essence of the stimulus-task presented to each of their *S*s. It was found that the stimulus-tasks presented differ essentially in terms of a right or left orientation of the stimulus-Gestalt. In view of the manner of stimulus-presentation, the high-low orientation of the stimulus-Gestalt was not problematic because *S*'s high-low orientation remained constant for all stimulus-presentations (*e.g.* $p =$ high, right orientation; $d =$ low, left orientation). By calling on Merleau-Ponty's studies of the relationship between the body and oriented space, it was found that the experimental results obtained by Natsoulas and Dubanoski could be accounted for by paying attention to the *relationship* between the stimulus-Gestalt orientation (right or left) and the body's right-left orientation. The *relationship* between the body orientation of *E* and that of *S* was found to be the basis for inferences concerning the locus of the perceiver.

University of California, Davis GEORGE D. YONGE

STIMULUS-VARIABILITY AND TWO MEANINGS OF MEANINGFULNESS *

We have shown that human *S*s are sensitive to the variability of random forms, and that there is an intermediate amount of variability (about 10 independent turns) which is consistently preferred by adults.[1] We have shown also that preference for stimuli varying in cognitive uncertainty is jointly determined by the number of independent stimulus-characteristics, rated meaningfulness, and *S*'s experience of variability. We used preference for random shapes as a measure of *S*'s response to variability. Finding that preferences shift systematically toward more stimulus-variability following experience of variability, and that the meaningfulness

* This work was supported by a grant from the Carnegie Corporation of New York.

[1] Harry Munsinger and William Kessen. Uncertainty, structure, and preference. *Psychol. Monogr.*, 78, 1964, 1-24. (No. 586)

of the high-variability forms affected preferences for the forms, we inferred that our *S*s had developed coding rules which permitted them to process greater variability.

The meaningfulness of random shapes has been determined and related to their variability several times,[2] but now the relation between variability and meaningfulness must be examined in more detail. The fact

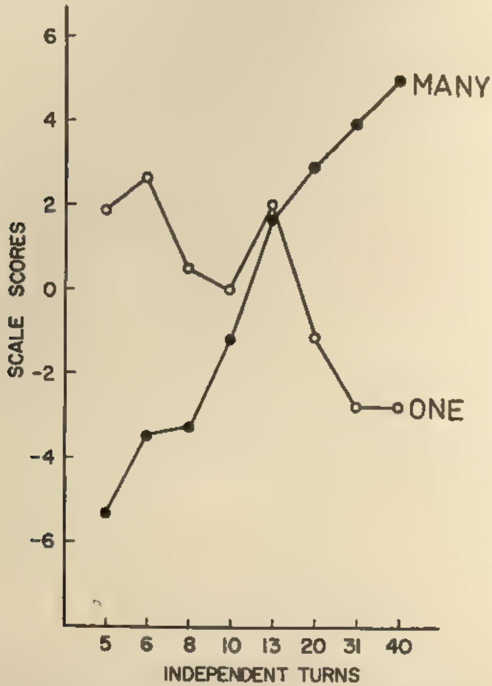

FIG. 1. MEANINGFULNESS AS A FUNCTION OF VARIABILITY UNDER TWO INSTRUCTIONS

that earlier investigators have found inverse and nonmonotonic relations between the variability of random forms and their meaningfulness suggests that two processes may be at work. On the basis of our own notions concerning the relation between meaningfulness and variability, we expect that, when *S* is asked to process or structure an entire form (*i.e.* when he is asked to give a single response which characterizes the form), he will find the task easier with forms of low variability than with

[2] See, for example, J. M. Vanderplas, and E. A. Garvin. The association value of random shapes, *J. exp. Psychol.*, 57, 1959, 147-154.

forms of high variability. When, instead, *S* is asked to report how many different things the form makes him think of, we expect the task to be easier with forms of high variability. To examine these two possible factors in the relation between variability and meaningfulness, we asked *Ss* to judge a series of random forms varying from 5-40 turns under two conditions of instruction. The *Ss* were 104 undergraduates drawn from classes in introductory psychology at the University of Illinois.

In a complete paired-comparisons design, 28 pairs of random shapes were presented. Seven examples at each of eight levels of variability were used. Each pair of stimuli was projected on the front wall of the room in which *Ss* were seated. The *Ss* of one group were instructed to pick the shape from each pair which reminded them more of "many things," while the *Ss* in the other group were instructed to pick the shape from each pair which reminded them more of "one thing," and these instructions were reversed during a second presentation of the stimuli.

A complete paired-comparisons scaling analysis was performed using a least-squares solution of Mosteller for a matrix *X* containing no vacant cells.[3] The results for the two conditions of instruction are presented in Fig. 1, which shows two different relations between variability and meaningfulness. The instructions to judge on the basis of the strength of a *primary* association ("one thing") produced an inverse relation between the number of turns in the figures and their judged meaningfulness, while the opposite relation was produced by instructions to judge on the basis of the total number of associations ("many things"). This finding strongly supports our notion of the effect of *S*'s ability to structure random shapes and the relation of this ability to the meaningfulness of shapes.[4] It is interesting to note that combination of the two curves shown in Fig. 1 (a result one might obtain with ambiguous instructions to judge "meaningfulness") produces a nonmonotonic function. The postulation of two ways of assigning meaningfulness to figures of differing variability thus permits the derivation of all forms of empirically obtained relations between variability and judged meaningfulness.

University of Illinois
Yale University

HARRY MUNSINGER
WILLIAM KESSEN

---

[3] Frederick Mosteller, Remarks on the method of paired comparisons. I. The least squares solution assuming equal standard deviations and equal correlations, *Psychometrika*, 16, 1951, 3-11; F. Mosteller and W. S. Torgerson, *Theory and Methods of Scaling*, 1958, 170-173.

[4] Munsinger and Kessen, Stimulus-variability and cognitive change, *Psychol. Rev.* (in press).

THE FIRST RECORDED OBSERVATION OF THE AFTER-EFFECT OF SEEN MOTION

After prolonged viewing of contours moving in one direction, a stationary contour appears to move in the opposite direction. This phenomenon has been termed the "waterfall illusion" or the motion after-effect (*MAE*). With the recent work on the *MAE* and its neurophysiological substrate it is of some interest to establish priority among those who observed and reported the after-effect.[1]

Wohlgemuth points out in his monograph on the *MAE* that the first recorded observation of any of the consequences of seen motion appears in the writings of Aristotle.[2] In *De Somniis,* Aristotle wrote: "when persons turn away from looking at objects in motion, *e.g.* rivers, and especially those which flow very rapidly, they find that the visual stimulations still present themselves, for things really at rest are then seen moving."[3] From the context of this passage and the argument pursued therein it may be concluded that Aristotle was describing an after-effect in the same direction as the river's motion and that he was not referring to the phenomenon which we call the *MAE.*

Most writers have credited R. A. Addams with the discovery of *MAE* in 1834.[4] Addams' description of the after-effect of watching a waterfall and of his success in producing *MAE* in his laboratory has been reprinted recently and this interesting report is now easily accessible.[5] Addams' report, however, was actually antedated by another, fourteen years earlier and we must instead credit the discovery of the *MAE* to the great Bohemian physiologist Johannes Evangelista Purkinje. Wohlgemuth correctly recognized Purkinje's priority, but overlooked the earliest of Purkinje's descriptions of the *MAE.*

In his paper, "Beiträge zur Nähren Kenntniss des Schwindels aus Heutognostischen Daten," published in 1820, Purkinje describes the

---

[1] R. W. Sekuler and Leo Ganz, Aftereffect of seen motion with a stabilized retinal image, *Science,* 139, 1963, 419-420; H. B. Barlow and R. M. Hill, Evidence for a physiological explanation of the waterfall phenomenon and figural after-effects, *Nature,* 200, 1963, 1345-1347.

[2] Wohlgemuth, On the after-effect of seen movement, *Brit. J. Psychol. Monogr. Suppl.,* 1, 1911, 1-117; W. A. Thalman, The after-effect of seen movement when the whole visual field is filled by a moving stimulus, this JOURNAL, 32, 1921, 429-441.

[3] Aristotle, De Somniis in W. D. Ross (ed.), *The Works of Aristotle,* 8, 1931, 459 b.

[4] R. A. Addams, An account of a peculiar optical phenomenon seen after having looked at a moving body, *London and Edinburgh Philosophical Magazine and Journal of Science,* 5, 1834, 373-374 (third series).

[5] W. N. Dember, *Visual Perception: The Nineteenth Century,* 1964, 80-81.

MAE as a kind of illusion which can be observed after extended viewing of a passing series of spatially distinct moving objects such as a long parade of riders, waves drifting by, or the spokes of a wheel revolving at an appropriate rate.[6] In that paper he also made brief reference to the explanation which he would offer some five years later, an explanation based on the effects of involuntary tracking eye-movements. Because the relevant material in the 1820 paper is so sparse, a translation of the somewhat longer discussion from the later paper, "Kenntniss des Sehens in subjectiver Hinsicht" more adequately conveys the flavor of Purkinje's writing. "Once I watched a procession of riders which lasted more than an hour. When the procession had passed, the houses which stood opposite me seemed to move in a direction opposite to that of the procession. Since the eye, while viewing the series of soldiers, had become accustomed to fixating every single one, it moved unconsciously in the same direction as the soldiers; this often repeated movement was habitual for a while, and it continued when the procession had passed. The eye still tended to fixate the stationary surroundings in the fashion to which it had become accustomed to fixating the things in motion. Therefore, it unconsciously slipped toward the accustomed direction and the surroundings seemed to slip away in the opposite direction. Of a similar kind is that illusory motion which the eye experiences after viewing quickly flowing water."[7] These observations are the same ones to which Purkinje alluded so briefly in the 1820 paper.

It is actually a little surprising that the parade would be sufficient to generate a *MAE*. We know that *MAE* is strongly dependent upon the angular velocity of the stimulus, *i.e.* the contours of inspection, and it is difficult to imagine that the angular velocity of the parade even approached the optimum angular velocity of about 4–6° of visual angle/second.[8] It is possible, however, that the extraordinary period of inspection, *i.e.* "more than an hour," may have compensated for the less than optimal angular velocity of the stimulus.[9]

---

[6] J. A. Purkinje, Beiträge zur Nähren Kenntnis des Schwindels aus Heutognostischen Daten, in K. J. Lhotak (ed.), *Purkyne's sebrane spisy* (*opera omnia*), 2, 1937, 23-24.

[7] Purkinje, Neue Beiträge zur Kenntniss des Sehens in subjectiver Hinsicht, in K. J. Lhotak (ed.), *ibid.*, 1, 1919, 92-93.

[8] Ragnar Granit, On inhibition in the after-effect of seen movement, *Brit. J. Psychol.*, 19, 1928, 147-157; Sekuler and Ganz, *op. cit.*, *Science*, 139, 1963, 419-420.

[9] Tõsaku Kinoshita, Ueber die Dauer des negativen Bewegungsnachbildes, *Z. f. Sinnesphysiol.*, 43, 1909, 434-442; W. A. Thalman, The after-effect of seen movement when the whole field is filled by a moving stimulus, this JOURNAL, 32, 1921, 429-441.

Purkinje's eye-movement explanation of the *MAE* has, of course, been discarded. In 1880, S. P. Thompson demonstrated that if opposite directions of movement were viewed simultaneously and monocularly, opposite direction of the *MAE* resulted.[10] An impossible kind of eye-movement is required to maintain an eye-movement explanation of opposite and simultaneous *MAE*s. Though Thompson's demonstration was compelling, it has been reinforced by the recent demonstration that the *MAE* can be generated under conditions in which eye-movements are compensated for optically, with stabilization of the retinal image.[11] Despite the fact that the details of Purkinje's explanation were wrong, the discovery of the *MAE* in 1820 must be added to his already impressive list of contributions to psychology.[12]

Northwestern University ROBERT W. SEKULER

ERRATA

In a paper by A. M. Hartman (Effect of reduction on the relationship between apparent size and distance, this JOURNAL, 77, 1964, 353-366), in Fig. 1, the trends of apparent size for Groups E_3 and E_4 are labeled inversely. In Fig. 2, the labels of the trends for apparent-distances are inverted for Groups E_3 and E_4, with the exception of the 371-cm. objective distance.

Attention is also invited to an error in R. C. Gonzalez and Bryan Shepp's study (The effects of endbox placement on subsequent performance in the runway with competing responses controlled) which appeared in the September issue, Volume 78, 1965, 445. In line 20, the 'connection' should be $s_{r_{eb}} \rightarrow r_G$ instead of $r_{eb} \rightarrow r_G$ *i.e.* a connection of stimulus and response instead of response and response.

K.M.D.

[10] S. P. Thompson, Optical illusions of motion, *Brain*, 3, 1880, 289-298.
[11] Sekuler and Ganz, *op. cit.*, *Science*, 139, 1963, 419-420.
[12] E. G. Boring, *Sensation and Perception in the History of Experimental Psychology*, 1942, 117 f., 177 ff., 277, 536, 588.

Abraham Aaron Roback: 1890–1965

A. A. Roback was in a certain sense a hermit scholar. During a period of more than 40 years, he lived successively in three roomy attic apartments in Cambridge. Books reached from floor to ceiling. The top of his desk was buried in correspondence and manuscripts. Part of the accumulation dealt with the publishing business of the Sci-Art Company which he managed almost singlehanded.

Much of his time was devoted to exchanges of letters, sometimes with little people in many lands, and sometimes with the truly great. *Freudiana* (1957) contains the text of letters he received from Freud, from Pavlov, Havelock Ellis, Romain Rolland, Bernard Shaw, and others. He compiled and edited two books on Albert Schweitzer, another of his longtime correspondents. Always he was partial to the personal and biographical, even anecdotal, approach. Dipping backward in time, he wrote *William James: His Marginalia, Personality and Contribution* (1942), *A History of American Psychology* (1952), and *History of Psychology and Psychiatry* (1962).

Perhaps the psychological work of most enduring value is *The Psychology of Character* (1927), setting forth both a history of the topic and a theory of character which has much in common with McDougall's mode of thought. Separately, he published his *Bibliography of Character and Personality* (1927).

While he kept in touch with many eminent psychologists and wrote widely on the fashions and schools of psychology, he was not a follower. Thus, in the early 1920s, he resisted the rise of behaviorism, publishing a polemical discussion, *Behaviorism and Psychology* (1923). Another deviation was his espousal of graphology, a topic he felt to be unjustly neglected by American psychologists. His interest in this field may have derived from his doctoral research (supervised by Münsterberg), and published as *The Interference of Will Impulses* (1917).

Folklore and etymology were likewise prime concerns. Beginning in 1928, he served as volunteer organizer of Harvard University's Yiddish collection, now a rich trove of 10,000 volumes. His own *Story of Yiddish Literature* appeared in 1940; *Peretz: Psychologist of Literature* in 1935; and a *Dictionary of International Slurs* in 1944.

To a list of 30 or more books one must add approximately 2000 articles, in Hebrew, Yiddish, German, as well as English—a prodigious bibliography indeed. He wrote rapidly and revised little. Although known in many lands for his psychological writing, it is probable that his fame is greater as an interpreter of Jewish, especially of Yiddish, culture.

Primarily, he was an author, only secondarily a teacher. He held brief appointments at Pittsburgh, Northeastern, Harvard, Massachusetts Institute of Technology, and Clark, but his longer assignments of teaching were on a part-time basis with the Massachusetts Department of University Extension and at Emerson College, in Boston.

Born on June 19, 1890, in Poland, of humble parentage, he grew up in Montreal, and graduated with high honors from McGill (B.A., 1912). Coming then to Harvard he received the M.A. degree in 1913, and the Ph.D. in 1917. He was a fellow of the American Psychological Association, likewise a governor of the Jewish Academy of Arts and Sciences, and a member of the Academic Council of YIVO. He died of a heart attack in Cambridge on June 5, 1965, two weeks before his 75th birthday.

Although he led a somewhat solitary life, a sincere loyalty and generosity marked his relationships with his friends. As a humanist he was consistently partial to philosophy, to literature and language, and to music. It was the humanistic spirit that fixed his outlook upon psychology. An appreciative essay written by Joseph Berger, entitled *The Destiny and Motivation of Dr. A. A. Roback,* was published by Schoenhof in 1957.

Harvard University GORDON W. ALLPORT

BOOK REVIEWS

Edited by T. A. RYAN, Cornell University

Strategy and Conscience. By ANATOL RAPOPORT. New York: Harper & Row, 1964. Pp. xxvii, 323. $6.95.

This book accurately reflects the breadth of competence and depth of concern of its author. Rapoport the mathematician describes the requirements which Game Theory must satisfy if it is to serve as a basis for rational decisions (Part I). Rapoport as social scientist and critic demonstrates how the use of Game Theory by the new and politically powerful profession of "strategic thinkers" fails to meet these requirements, despite their claims to rationality and objectivity (Part II). And Rapoport the humanist with a conscience—historian, semanticist and once concert artist—explores the contrasts between American and Soviet ideologies and practices in a search for perspective on foreign policy (Part III).

Three modes of decision theory are distinguished: A formal, mathematical theory requires no data, by definition, but is a deductive system for deriving logically necessary conclusions from given assumptions. A descriptive, behavioral theory makes use of whatever data are available or can be generated in experiments; it seeks to discover the criteria which guide real people's decisions. A prescriptive, normative theory requires specific, valid data about the real world, without which the theory can yield no prescriptions for action.

As games become more complex, the requirements which must be met for rationality increase in number and subtlety. Where outcomes of actions are certain, rationality requires only that preferences be consistent, transitive and coupled directly with appropriate acts—yet even here humans are not always rational on the descriptive level. Where outcomes are risky, both utilities and probabilities of the various outcomes must be assigned; the former assumes averaging "over the long run" and the latter ordinarily assumes observed frequencies—"subjective probabilities" come down to degrees of belief. In so-called Games Against Nature, a player need not take into account the motives of the opponent, although even here players often adopt optimistic (maximax) or pessimistic (minimax) strategies.

Real games are defined by having two or more players with at least partially conflicting interests, with some range of choices called "strategies," where the outcomes or "payoffs" are determined once each player has selected a strategy. In zero-sum games (where A's gain is B's loss and vice versa) assumption of rationality of the opponent is inherent, according to Rapoport. But in non-zero-sum games, like the Prisoner's Dilemma or arming vs. disarming in an international conflict, assumptions about the motives of the opponent, about just what is really "rational" and even about the role of "conscience" become crucial—and, of course, it is such mixed-motive games which most closely approximate real-life situations.

Rapoport's critique of "strategic thinking" as it is being applied today in the

determination of national policy is, in this reviewer's opinion, devastating. Necessarily at the prescriptive level—for example, as to whether and how much to invest in Civil Defense—such thinking is shown to be based upon subjective estimates of utilities (What price in millions of American lives are we willing to pay to punish Soviet aggression?) and probabilities (a difference between 0.01 and 0.001 in the likelihood an opponent will be as stubborn as we are in a crisis makes a tremendous difference in "calculated risk" yet it cannot be determined). Faced with the almost insurmountable difficulty of assigning utilities of outcomes to opponents, without which there can be no rigorous prescription in a non-zero-sum game, our strategists are under pressure to convert real world conflicts into zero-sum games—where the only motive of the opponent is aggressive self-interest. Lacking any objective basis for assigning utilities and probabilities, prescriptive game theory requires a firm descriptive base in decision-making at the behavioral level—yet, according to Rapoport, ". . . the paucity of psychological knowledge among strategic thinkers is appalling . . ." (p. 123). In sum, according to the author, such recommendations for policy decisions, as if based on rational analysis derived from "calculated" risks, are fraudulent in effect if not in intent.

Readers of this JOURNAL, as experimental psychologists, will be particularly interested in Rapapart's description of his own simulations involving two-person, non-zero-sum games of the Prisoner's Dilemma type. Characteristic of the author's humility, this description is placed in Part II, along with his criticisms of mis-use of Game Theory, and ends with a warning about the limited generality of the results. Nevertheless, the experimentation is extensive, systematic and full of intriguing implications. It is an attempt to build some of the psychological under-pinnings for a more realistic prescriptive theory. For example, we find that knowledge of the payoff matrix increases cooperative choices; correlation between the patterns of choices of players increases over a long sequence of plays—a kind of reciprocal "shaping" of behavior? In the interest of further realism, this reviewer would like to see successive moves introduced (simultaneous moves are standard practice), the possibility of graduation in risk at the option of the player (all-or-nothing payoffs are the rule), and analysis of data in terms of larger units (Rapoport demonstrates the extraordinary increase in strategic complexity when sequences of even two plays are considered, p. 53)—but such modifications might put the Game beyond rigorous mathematical treatment.

Rapoport's plea for consideration of conscience in decision-making is more than emotionalism—although there is no question about the depth of his feeling. In his own experiments and those of others he can point to the role of variables like "trust," "greed," "repentance," and "forgiveness" in the play of non-zero-sum games. In other words, "conscience," no matter how it may elude anything other than operational definition, is a significant factor in decision-making. But can there be an effective dialogue between strategic thinkers and conscience-driven thinkers? Rapoport thinks not. Whereas the conscience-driven thinker must emphasize effects of actions upon actors (rather than upon "nature"), the determination of utilities (rather than taking them as given), and the self-fulfilling propensities of strategic thinking itself, the strategic thinker minimizes these factors. In reaching this pessimistic conclusion, Rapoport seems to attribute lack of conscience to strategic thinkers as well as lack of humility about their methods. It may well be that the

major impact of this book will be to create conscience out of humility or humility out of conscience, depending on which the reader possesses in greater measure.

Strategy and Conscience concludes with an historical analysis of the development of American and Soviet ideologies (and practices) which is fresh, sensitive and full of psychological insights. This analysis leads Rapoport to the conclusion that each society has something which the other needs. American society has political freedom from compulsive conceptualizations and conformities; Soviet society has a social conscience that is appropriate to a complex and crowded civilization. Anatol Rapoport recommends a kind of "ideological disarmament," in which each society asks what can be learned from the other for one's own benefit rather than castigating the other for what it does not have.

University of Illinois — CHARLES E. OSGOOD

Social Change and Prejudice. By BRUNO BETTELHEIM and MORRIS JANOWITZ London, Free Press of Glencoe, 1964. Pp. xi + 337. $5.95.

This book contains a reissue of the authors' earlier work, *Dynamics of Prejudice,* 1950, and what is intended to be a reassessment of the conclusions of that study based upon data and theory of the past fifteen years. A reassessment was certainly indicated; the original findings were descriptive in nature and based principally upon the self-reports of a very restricted subject sample during a unique period of history, while the major conclusions were in terms of cause-effect relationships and broadly generalized. The extent to which the needed reassessment is accomplished is, however, a point of question.

The authors emerge with the general conclusion that their original formulations have been well substantiated, but the current data are no less tenuous, and interpretations remain less than conservative and appear, on occasion, to spring more from the authors' theoretical predilections than the facts. For example, they attribute an increase in anti-semitism between 1940-1946 to the tensions of the war, and a subsequent decrease to social change and economic security concommitant with industrialization and the availability of the Soviet Union as an object of displaced aggression. One has the uncomfortable feeling, however, that if the trend were in the opposite direction, it could have been interpreted just as plausibly within the same framework. The war years constituted a time of rapid industrial expansion and individual economic security, and objects for hostility were certainly present. For the subsequent period, tensions could have been ascribed to post-war adjustment and cold-war in a nuclear age.

Further in the text the authors consider changes in age-groupings as a factor in the decrease of prejudice during 1946-1959, and state that "examination of changes in gross age-composition during this period shows them to be unimportant factors in the trends in prejudice. The age structure is still relatively stable, decade by decade . . ." (p. 17). No indication is given of how units of age-change may be related to units of prejudice-change, nor is such an analysis possible, and the degree to which one variable may account for the other appears to be entirely a matter of conjecture. In commenting upon the relationship of education to prejudice, the conclusion is offered that "on the basis of some 25 national sample surveys since 1945 . . . lower levels of prejudice among the better educated seem to involve the social experience of education specifically and not merely the sociological origins of the educated" (p. 18). The findings referred to indicate that

better educated people are less likely to hold a "provincial outlook" and "primitive misconceptions" with regard to minority groups. These differences may reflect a direct influence of the education process which causes a decline in prejudice, but they serve also in these studies as a measure of prejudice, and you cannot employ the same operational definition for the antecedent and consequent variables and arrive at a meaningful conclusion.

The authors discuss seven studies dealing with the mobility-prejudice hypotheses of *Dynamics of Prejudice, i.e.* downward mobility is associated with increased prejudice and upward mobility with increased tolerance, and these may be the most salient of the data presented for a reassessment of the original findings. For this reason, it is particularly unfortunate that information about these studies is not more complete. In one instance, the authors offer solely a conclusion based on "special tabulations" of data presented in a book which is cited, but without page references. Perusal of the book indicates that the location of the data and the general nature of the special tabulations may be conjectured, but are not readily apparent. We are told that for tests of the downward mobility-prejudice hypotheses, the mobile groups were compared to the lower status groups they entered, which is essential to ascertain the effects of mobility, *per se.* We are not, however, given similar information regarding the upward mobility-prejudice relationships. In two cases reference is made to different findings with different subject samples, but no indication is given as to the nature of these sample differences.

Nevertheless, the book does demonstrate that there is substantial support for the downward-mobility-prejudice relationship and at least some support for the converse hypothesis. It must be noted, however, that there is still no evidence that mobility *causes* attitude change, which is the assumption underlying the authors' extensive discussion of implications of the findings here and in the earlier book. There are any number of potential latent factors which may account for the mobility-prejudice relationships, and it is surprising that no one directly concerned with the question has attempted the longitudinal study that will resolve it.

It must be expected that an endeavor of the magnitude of a sociopsychological theory of minority-group prejudice will entail shortcomings. Despite those of *Social Change and Prejudice,* there is a good deal of information in this book, and the authors' interpretations, if controversial, are generally provocative. It will be of more than passing interest to the student of prejudice and should generate some of the better research of the next decade.

State University of New York at Buffalo IRWIN SILVERMAN

Human Senses and Perception. By G. M. WYBURN, R. W. PICKFORD, and R. J. HIRST. Toronto, University of Toronto Press, 1964. Pp. xii, 340. $7.00.

According to the authors, the purpose of the book is to provide in a single source "information that a biologist, psychologist, or philosopher might wish to have concerning those aspects of the subject [senses and perception] outside his special interest" (p. vii).

The treatment consists of three parts. For the sake of simplicity we shall review critically each part and thereafter consider the value of the work as a whole.

Part I is devoted to a rather cursory presentation of the sense organs. The illustrative material is not original and often fails to serve usefully because it is either

ignored in the text or inadequately treated. Some figures are not labelled at all and are consequently very confusing (*e.g.*, p. 3), while others are too complex for the level of the text (*e.g.*, p. 73). There are, in addition, some unfortunate errors and ambiguities. The authors state that the magnitude of a neural impulse in any given neuron is always constant. They ignore completely the role of the relative refractory period, accommodation, fatigue, and metabolic support. Moreover, the constancy of impulse magnitude is confused with the all-or-none law. Elsewhere, "the middle ear communicates with the back of the nose" (p. 51), the authors refer to the "tympanic membrane of the eardrum" (p. 51), and alpha rhythm is reported as having a frequency of six per minute (p. 109).

For an elementary exposition the authors resort frequently to technical terms, such as *cytoarchitecture, medial lemniscal system,* and *centrifugal,* without pity for the particular reader for whom the book was purposely written.

Part II deals with perception. The illustrative material is handled successfully, and, although discursive, the style of writing is noticeably superior to Part I. The theoretical bias favors Gestalt principles in perception. A particularly thorough discussion of the perception of time, movement, and causality adds a freshness to what might otherwise have been just a conventional presentation of perceptual topics.

The authors begin Part II with polite criticism of behaviorism, but they seem not to appreciate the metamorphosis of behaviorism since the days of Watson. They try unsuccessfully to distinguish perceptual research in animals and man, claiming that, "all studies of perception in animals are . . . inferential" (p. 136). They go on to question the suitability of the natural sciences to explain perception.

There are several minor sources of confusion in Part II. The terms *optic disc* and *blind spot* are used indiscriminately. Dichromats are said to see only two colors (p. 205). In discussing the perception of size and shape the authors fail to make plain what is meant by "object" vs. "stimulus" attitude on the part of the observer (p. 157-159).

There is no discussion of the psychophysical methods or their relevance to perceptual data.

Part III purports to deal with philosophical issues relating to perception which are of interest to the biologist and psychologist. Here the authors search for metaphysical truth, wrestle with a suitable definition of perception which avoids the role of the senses and the cortex, and finally after 100 pages make a plea for a fresh approach.

The historical perspective given to some of the concepts of perception is useful. The section on primary and secondary qualities is especially lucid. Yet, Part III adds very little to the substance of the book while adding unnecessarily to its length.

The book considered as a whole is of limited value because its purpose is too ambitious to be met adequately. The presentation ranges in level from the very simple to the highly sophisticated, and one is often left with a feeling of despair that so many topics were discussed so superficially. The omission of psychophysics, efferent modulation in sensory channels, and central brain processes in perception makes the book outdated before its time.

W. L. Gulick

University of Delaware

Motivation and Emotion. By E. J. MURRAY. In R. S. LAZARUS (ed.) *Foundations of Modern Psychology Series.* Englewood Cliffs, New Jersey, Prentice-Hall, 1964. Pp. ix, 118. $1.50.

This paper-back book is one of fourteen in the *Foundations of Modern Psychology Series.* The aim of the series is to provide the instructor with a group of short volumes, each a self-contained book dealing with the special issues, methods, and content of some basic area of psychology. With such a series available the "teacher can choose the subjects he wants to emphasize and present them in the order he desires. And without necessarily sacrificing breadth, he can provide the student with a much fuller treatment of individual areas at the introductory level than is normally possible." (p. v) The book under review is a sample of the series.

Murray's *Motivation and Emotion* gives a broad, up to date coverage of topics ordinarily included in textbook discussions of this complex and disorganized area of psychology. The author considers the entire spectrum of behavioral determinants: hunger, sex, fear, anger, curiosity, affection, need for achievement, affiliation, and other motivations. Murray examines the dominant theories of instinct, drive, homeostasis, hedonism, cognitive theory, intrinsic motivation, and special theories including Festinger's dissonance, White's effectance motivation, Woodworth's behavior-primacy doctrine, Maslow's self-actualization. In discussing the development of social motivation Murray considers Freud's view of development through sublimation. He gives an excellent survey of contemporary views of sexual motivation with a critique of Freudian theories. He considers G. Allport's concept of functional autonomy, N. Miller's work on acquired motivation, the studies of McClelland, Atkinson, and others, on achievement motivation and affiliation. Illustrative materials are drawn from diverse fields—human and animal experiments, clinical observations, anthropological field studies, and related sources. The coverage of the book is thus exceedingly broad.

Murray describes his approach as pluralistic and integrative but there is nothing original about his views. He defines a motive as "an internal factor that arouses, directs, and integrates a person's behavior" (p. 7), and comments that motives are inferred to explain behavior and not directly observed. Murray's definition of emotion is similar to that of Magda B. Arnold. Fear and anger are treated as motives; the disruptive aspect of strong emotions is considered as something secondary to emotional motivation rather than as a separate aspect of behavior.

The book is clearly written, readable, with twelve figures which the undergraduate student can easily understand. It is definitely an introductory text, not an advanced work.

Motivation and Emotion is up to date. At the close of the book 32 selected readings are listed. Of these, 20 were published during 1960-1965; 11 during 1952-1959; and 1 (the first Kinsey report) in 1948. The selected readings reflect the current explosion of studies in this complex and rapidly expanding area of psychology. If the explosion continues, the book will need revision in a few years.

How important is this series of books to the teacher of psychology? I think the answer depends upon the aim and plan for an introductory course. If the teacher wanted to emphasize a few topics (for example, perception, motivation, learning, physiological psychology) and ignore others (for example, language

and thought, personality and adjustment, social psychology, tests and measurements), he could select several of these books, according to his interests, and build a beginning course around them. If he wanted to survey the entire field of modern psychology, he might prefer to use a single text such as Hilgard, Munn, or some other excellent book selected from the plethora now available. Should the introductory course be built around a series of paper-back books or around a single text? The teacher will decide.

Claremont, California PAUL THOMAS YOUNG

Leaders, Groups, and Influence. By E. P. HOLLANDER, New York, Oxford University Press, 1964. Pp. xiv. 256. $5.00.

A more appropriate title for this book would have been "Selected Readings from the Publications of E. P. Hollander." The volume consists of 20 chapters, three of which have been written especially for this book. The remaining chapters are adaptations from previously published papers by Hollander and his collaborators. The cohesion among the 17 reprinted chapters is small and certainly no greater than that obtained in most books of readings organized around some specific research topic. The author attempts to provide cohesion by overviews in the introduction and in the final chapter, by frequent cross-references to other chapters, and by introductory summaries to the various sections of the book. But these efforts are not completely successful.

Hollander is a social psychologist of considerable talents. His writings are solid, useful, and important. Thus, from some points of view it was valuable to bring them together in a single volume. From another point of view, however, it may be said that this particular presentation will not make an important contribution to the field, over and above the contributions which the author has already made when he originally published the particular papers. The present reviewer feels that the book lacks a clear focus. Furthermore, he thinks that the organization of the key ideas into a 24-page *Psychological Review* article would have made a greater contribution than the present 238 pages. In this day and age, when we are flooded with publications, it is desirable for authors to present their points of view as succinctly as is possible.

The essence of Hollander's point of view was presented in his 1958 article in the *Psychological Review.* In an attempt to explain how leaders can be greater "conformers" to group norms, while also being greater "innovators," be postulated the existence of "idiosyncracy credit." The more effective leaders make it possible for their followers to reach satisfactions associated with group success, and are allowed, therefore, a certain amount of "credit" by their followers. Idiosyncarcy credit permits the leader to engage in behaviors which may be perceived by the group to deviate from norms and yet not experience sanctions from the group members. Thus, a leader with much idiosyncracy credit can engage in innovations and have influence with group members, while a leader with little credit can do neither.

The present volume brings together the articles which are related to the elaboration of this point of view, and the articles which report on the empirical evidence in support of this conception. Since the volume of these materials was clearly insufficient for a book, the author has assembled additional papers related to the general theme of "leadership" under a single cover. These additional papers in-

clude Hollander's most valuable studies on the peer-nomination technique, his studies of leadership, followership and friendship, group consensus, and group attraction.

A detailed review of each of these articles will not be undertaken, since it would be even more unfocused than the volume under review. A brief comment on the main thesis is possible.

It is the present reviewer's impression that while the thesis is plausible and promising, it is not as yet fully developed. It seems likely that in a few years the author will have enough empirical evidence to publish a book on his major thesis. The present volume, however, is a premature anticipation of such a publication.

University of Illinois HARRY C. TRIANDIS

A Social Psychology of Group Processes for Decision-Making. By B. E. COLLINS and H. GUETZKOW. New York: Wiley, 1964. Pp. x, 254. $6.25.

The purpose of this book is two-fold: To bring together the vast literature which is relevant to decision-making in groups and, at the same time, to provide the decision-maker with a set of principles which will help him understand the processes involved in group-functioning. Both of these objectives are accomplished quite successfully.

In spite of its small size, the book is a rich source of information about small-group research. The writing is lucid and contains a minimum of technical jargon, which should make this book appealing to a wide range of readers. Anyone interested in the current status of our knowledge of group-processes will find this book attractive reading; there is something in it for everyone: layman, student, and professional worker.

The chapters cover the following areas: group vs. individual performance, task and interpersonal factors in group-productivity, obstacles to effective interpersonal relations in a group, direct and indirect sources of power in groups, power and group-member behavior, communication and interaction, participant satisfaction in a conference, and leadership. The authors use a propositional framework as a means of expressing the generalizations they have drawn from the literature, and present the evidence to support each proposition—sometimes using a whole chapter to provide full documentation. Of necessity, the authors have been selective in their reviews, but they have not ignored contradictory evidence. Where contradictions exist, the authors attempt resolutions which, although highly speculative, suggest promising directions for further research.

Because of the inductive approach taken in this book, theories of decision-making, as they might apply in groups, do not receive extensive treatment. In some cases, the propositions seem to imply causative relationships, but are stated so generally that the direction of causation is not clear; in other cases, propositions in different chapters tend to overlap with each other because the same basic phenomena are being treated from different points of view. This reviewer, at least, felt the need for a final summary chapter in which the numerous propositions were brought together in some type of organized general framework. In addition, some readers may find some of the descriptions of supporting studies too sketchy to provide them with an adequate understanding of the findings.

In spite of some shortcomings, Collins and Guetzkow have done us a useful service in providing such a compact guide to an important area of social psychology.

Indiana University SEYMOUR M. BERGER

Role Development and Interpersonal Competence. By DAVID MOMENT and ABRAHAM ZALEZNIK. Boston, Harvard University, Division of Research, Graduate School of Business Administration, 1963. Pp. xvi, 346. $6.00

This book reports a research study which had as its general purpose an investigation of how the behavior of the individual in a problem-solving group is related to his personal developmental history. The report is clearly written and well organized with approximately one half of the pages devoted to technical appendices.

The 52 members of four experimental groups completed a set of background questionaries prior to viewing a filmed case-problem and participating in a group-discussion. After the discussion they completed other questionaries to report their reactions to each other. On the basis of an analysis of information obtained from the later questionaries, the 52 *Ss* were designated as either Stars (15 members), Technical Specialists (10 members), Social Specialists (12 members), and Underchosen (15 members). Other information was then analyzed for association with these four "role-types."

The authors do an unusually thorough and competent analysis and interpretation of their data. In the process they develop a number of quite reasonable conclusions about differences in behavioral patterns, orientations toward tasks, motivations, and defensive systems that relate to the basic fourfold typology. A major problem in evaluation of their conclusion stems, however, from very serious questions which must be raised concerning the adequacy of the data and of the rigor that was maintained in the experimental situation within which they were obtained. It is doubtful that effort expended in analysis and interpretation is warranted if one judges from the point of view of demanding even minimal standards of careful experimental work. It would seem prudent therefore to regard the conclusion of the study to be as yet unverified hypotheses. These may have value to other researchers who might wish to pursue them further.

Educational Testing Service JOHN K. HEMPHILL

Human Nature in Politics: The Dynamics of Political Behavior. By JAMES C. DAVIES. New York, John Wiley and Sons, 1963. Pp. xii + 403. $7.50.

This social psychology of political behavior brings considerable acquaintance with the recent literature to bear upon the individual (as distinguished from institutional) aspects of political phenomena. The book is written in a breezy journalistic style that mingles anecdotes with research references. Apart from the title borrowed from Graham Wallas' armchair classic, its claims and manner are modest, though its intended scope is not. The psychologist may be put off by the author's penchant for glib parlor analyses of political motivation; in general, however, psychological sources are represented fairly, and Davies seems sensible enough in the applications that he makes of them to politics. Some of these applications struck me as particularly felicitous, as when he calls upon Brozek's hungry conscientious objectors and Lazarsfeld's *Arbeitslosen* to support a view of apathy as the true political condition of the Hobbesian "state of nature."

Since the book is too loosely constructed to make an original systematic contribution, the psychologist will be interested mainly in the aspects of psychology that are being relayed here to students of political science. Maslow's motivational hierarchy once again does heavy duty (it seems to be a favored export product); so also does New Look perceptual functionalism. Problems of socialization, group relations, and leadership tend to be approached quite directly in the political context, rather than by way of generalization from psychological research. Only the surface of the literature on persuasive communication is mined. Reference to group-theory and role-analysis are essentially ignored.

The student will be likely to enjoy and will probably benefit from this wide-ranging and liberal-spirited book, but the difficult systematic job of organizing from accumulated research and theory a relevant psychological foundation for political science remains largely to be done.

Center for Advanced Study in the Behavioral Sciences M. BREWSTER SMITH

Perception and the External World. Edited by R. J. HIRST. New York, The Macmillan Company, 1965. Pp. v, 310. $1.95 (paper).

This paperback, by a philosopher at Glasgow, is an edited set of readings on theories of perception. By careful selection, by annotation, and by providing a lucid introduction, this book becomes superior to most collections, and is to be recommended. It provides the only short general survey in existence of what philosophers have been writing in this century about the nature of sense-perception.

After the academic separation of the two disciplines psychologists have tended to quit reading philosophy, but the philosophers have gone right on with their interest in sensory and perceptual experience. Sometimes they even read experimental psychology. They never gave up their right to observe and introspect, even if they relinquished experimentation, and they have worked hard at the clarification of muddles. Theories of perception are full of the latter. Many experimental psychologists have recently been coming up with what they thought were new theories of perception but which were in fact old ones. They make assumptions about epistemology without knowing it. This book would be an antidote for such tendencies.

There are 23 readings here, some of them being difficult, as only philosophers can be difficult, but some of them being easy and forthright, like the selection from Gilbert Ryle. Only one of them is by a psychologist, Koffka. The book is aimed at students of philosophy. But the editor has classified them (and provided further references) so that the various theoretical positions of idealism, phenomenalism, and realism become fairly clear. The prolonged arguments about sense-data and the various efforts to justify some form of realism become intelligible. The book provides a quick way for the experimental students of perception to catch up on the philosophical analysis of perception.

Cornell University J. J. GIBSON

Work and Motivation. By VICTOR H. VROOM. New York, John Wiley and Sons, 1964. Pp. ix, 331. $6.95.

This is a sensible and informative critical summary of research in the area of motivation at work, which furnishes a long-needed replacement for Viteles' 1953

summary. It will probably be widely used as a survey at both undergraduate and graduate levels.

Vroom makes explicit an essentially additive model relating the valences and expectancies of alternative actions to the worker's choice of behavior, and applies the model to such problems as differences in effort to perform on the job. The model seems over-simplified, in view of the curvilinear relationships to be expected, but it is at least a clear first approximation.

Cornell University PATRICIA C. SMITH

The Frontiers of Management Psychology. Edited by GEORGE FISK. New York, Harper & Row, 1964. Pp. xi, 301. $6.00

Another symposium of uneven quality and heterogeneous content; this book includes presentations by economists, sociologists, business administrators, psychiatrists, and psychologists. Topics are grouped under "Values of Executives," Conflict in Decision Making," "Industrial Psychology," Applications in Functional Fields," and "After Work." Psychologists should look at this book if only for Gomberg's version of what really happened in Coch and French's study of group decision (his union was involved at that time), for Shostak's clear reminder of the indifference of most psychologists and most unions to each other, for Fisk's own tabulations of leisure time activities of persons in various categories of employment. There are also some very bad sections, including a paper on mental illness, barely suitable for a junior high-school course in mental health, and an object-lesson in bad methodology, validating the Rorschach for executive selection.

Cornell University PATRICIA C. SMITH

BOOKS RECEIVED

(The books listed here have not as yet been noted in our pages. Listing here does not, however, preclude their later review.)

ARNHEIM, RUDOLF, *Art and Visual Perception: A Psychology of the Creative Eye.* Berkeley, Univ. of California Press, 1965. Pp. xi, 485. $3.95.

BAILEY, PERCIVAL, *Sigmund The Unserene: A Tragedy in Three Acts.* Springfield, Ill., Charles C. Thomas, 1965. Pp. xvii, 127. $5.75.

BARRAL, MARY ROSE, *Merleau-Ponty: The Role of the Body-Subject in Interpersonal Relations.* Pittsburgh, Duquesne Univ. Press, 1965. Pp. xi, 297. $6.75.

BARRY RUTH, and WOLF, BEVERLY, *Motives, Values, and Realities: A Framework for Counseling.* New York, Teachers College Press, 1965. Pp. x, 278. $4.95.

BERNE, ERIC, *A Layman's Guide to Psychiatry and Psychoanalysis.* New York, Grove Press, 1962. (First published 1947.) Pp. xxi, 320. $.95 (paper).

CERDA, ENRIQUE, *Una Psicologia de Hoy.* Barcelona, Editorial Herder, 1965. Pp. xvi, 709. (paper).

CONRAD, JOHN, *Crime and its Correction: An International Survey of Attitudes and Practices.* Berkeley, Univ. of California Press, 1965. Pp. 312. $7.50.

CUTLIP, SCOTT M. (Ed.), *A Public Relations Bibliography.* Madison and Milwaukee, Univ. of Wisconsin Press, 1965. Pp. xiv, 305. $7.50.

DAVIDSON, HENRY A., *Forensic Psychiatry.* New York, Ronald Press, 1965. Pp. x, 473. $10.00.

DIENES, Z. P., and JEEVES, M. A. *Thinking in Structures.* Vol. I. London, Hutchinson Educational Ltd., 1965. Pp. 125.

FOWLER, HARRY, *Curiosity and Exploratory Behavior.* New York, Macmillan, 1965. Pp. viii, 216. $1.95 (paper).

FREEMAN, LINTON C., *Elementary Applied Statistics for Students in Behavioral Science.* New York, London, and Sydney, John Wiley & Sons, 1965. Pp. viii, 298. $6.95.

GEBHARD, PAUL H., GAGNON, JOHN H., POMEROY, WARDELL B., and CHRISTENSON, CORNELIA V., *Sex Offenders: An Analysis of Types.* New York, Harper & Row, 1965. Pp. xxxiv, 923. $12.50.

GELDARD, FRANK A. (Ed.), *Communication Processes.* Vol. IV. New York, Macmillan, 1965. Pp. x, 299. $14.00.

GUILFORD, J. P. *Fundamental Statistics in Psychology and Education.* Rev. Ed. New York, McGraw-Hill, 1965. Pp. 605. $8.50.

HAMACHEK, DON E. (Ed.), *The Self in Growth, Teaching, and Learning: Selected Readings.* Englewood Cliffs, N. J., Prentice-Hall, 1965. Pp. xi, 576. $7.50 (paper).

HARTLEY, RUTH E., and HARTLEY, EUGENE L. (Eds.), *Readings in Psychology:* Third Edition. New York, Thomas Y. Crowell, 1965. Pp. x, 612. $4.25 (paper).

HATCH, RAYMOND N., and STEFFLRE, BUFORD, *Administration of Guidance Services.* Second Edition. Englewood Cliffs, N.J., Prentice-Hall, 1965. Pp. xi, 466. $8.95.

HERRNSTEIN, RICHARD J., and BORING, EDWIN G. (Eds.), *A Source Book in the History of Psychology.* Cambridge, Harvard Univ. Press, 1965. Pp. xvii, 636. $12.50.

LEVITAS, G. B. (Ed.), *The World of Psychoanalysis.* New York, George Braziller, Inc., 1965. Pp. xxi, 1113. $12.50.

LINDZEY, GARDNER, and HALL, CALVIN S. (Eds.), *Theories of Personality: Primary Sources and Research.* New York, John Wiley & Sons, 1965. Pp. xiii, 543. $6.95.

LUIJPEN, WILLIAM A., *Phenomenology and Metaphysics.* Pittsburgh, Duquesne Univ. Press, 1965. Pp. xii, 202. $5.50.

LYONS, JOSEPH, *A Primer of Experimental Psychology.* New York, Harper & Row, 1965. Pp. xii, 322. $3.50 (paper).

MARKS, JOHN, and PARE, C.M.B. (Eds.), *The Scientific Basis of Drug Therapy in Psychiatry.* New York, Macmillan, 1965. Pp. xiii, 217. $10.50.

ORR, DOUGLASS W., *Professional Counseling on Human Behavior: Its Principles and Practices.* New York, Franklin Watts, 1965. Pp. xiii, 284. $5.95.

PINCKNEY, EDWARD R., and PINCKNEY, CATHEY, *The Fallacy of Freud and Psychoanalysis.* Englewood Cliffs, N.J., Prentice-Hall, 1965. Pp. v, 186. $4.95.

PROKASY, WILLIAM F. (Ed.), *Classical Conditioning: A Symposium.* New York, Appleton-Century-Crofts, 1965. Pp. ix, 421. $7.50.

REGER, ROGER, *School Psychology.* Springfield, Ill., Charles C. Thomas, 1965. Pp. xv, 213. $7.75.

SANFORD, FILLMORE H., *Psychology: A Scientific Study of Man.* Second Edition. Belmont, Wadsworth Publishing Co., 1965. Pp. 627. $7.95.

SCHEIN, EDGAR H., and BENNIS, WARREN G., *Personal and Organizational Change Through Group Methods: The Laboratory Approach.* New York, John Wiley & Sons, 1965. Pp. x, 376. $8.25.

TALLAND, GEORGE A., *Deranged Memory: A Psychonomic Study of the Amnesic Syndrome.* New York, Academic Press, 1965. Pp. viii, 356. $9.50.

THOMAS, CAROLINE BEDELL, ROSS, DONALD C., and FREED, ELLEN S., *An Index of Responses to the Group Rorschach Test: Studies on the Psychological Characteristics of Medical Students—II.* Baltimore, Md., The John Hopkins Press, 1965. Pp. xliv, 502. $15.00.

ULLMANN, LEONARD P., and KRASNER, LEONARD (Eds.), *Case Studies in Behavior Modification.* New York, Holt, Rinehart & Winston, 1965. Pp. xii, 401. $7.50.

VERNON, DAVID T. A., FOLEY, JEANNE M., SIPOWICZ, RAYMOND R., and SCHULMAN, JEROME L., *The Psychological Responses of Children to Hospitalization and Illness.* Springfield, Ill., Charles C. Thomas, 1965. Pp. vi, 192. $8.00.

ZEIGARNIK, BLYUMA VUL'FOVNA (Translated by Basil Haigh), *The Pathology of Thinking.* New York, Consultants Bureau Enterprises, 1965. Pp. xv, 211. $12.50.

ZUBIN, JOSEPH; ERON, LEONARD D., and SCHUMER, FLORENCE, *An Experimental Approach to Projective Techniques.* New York, John Wiley & Sons, 1965. Pp. xix, 645. $13.50.

INDEX

CHARLES VALLEY BROOK, University of Texas

AUTHORS

(The names of authors of original articles are printed in CAPITALS and SMALL CAPITALS; of authors of books reviewed in roman, and of reviewers in *italics*.)

SUBJECTS

(References in *italic* figures are to reviews.)